

**ВАЗОРАТИ МАОРИФ ВА ИЛМИ ҶУМҲУРИИ ТОҶИКИСТОН
ДОНИШГОҲИ ДАВЛАТИИ ҲУҚУҚ, БИЗНЕС ВА
СИЁСАТИ ТОҶИКИСТОН**

ВБД: 811.21; 004.9

Бо ҳуқуқи дастнавис



АНВАРЗОДА АКМАЛ АНВАР

**МАСЪАЛАҲОИ ШИНОХТАГИРИИ ФАРМОНҲОИ ОВОЗӢ ВА
ТАТБИҚИ ОНҲО ДАР МОДЕЛИ КОМПЮТЕРӢ**

АВТОРЕФЕРАТИ

диссертатсия барои дарёфти дараҷаи илмии доктори фалсафа (PhD) - доктор аз рӯйи ихтисоси 6D070400 – Техникаи ҳисоббарор ва таъминоти барномавӣ (6D070402 – Мошинҳои ҳисоббарор, комплексҳо ва шабакаҳои компютерӣ)

Хуҷанд – 2025

Диссертатсия дар кафедраи технологияҳои иттилоотӣ коммуникатсионӣ
ва барномарезии Донишгоҳи давлатии ҳуқуқ, бизнес
ва сиёсати Тоҷикистон иҷро шудааст

Рохбари илмӣ:

Худойбердиев Хуршед Атохонович -
доктори илмҳои техники, дотсенти кафедраи
барномарезӣ ва низомҳои иттилоотии
донишкадаи политехникии донишгоҳи
техникии Тоҷикистон ба номи академик
М.С. Осимӣ дар шаҳри Хучанд

Муқарризони расмӣ:

Рахимов Нодир Одилевич –
доктори илмҳои техники, дотсенти кафедраи
таъмини барномавии технологияҳои
иттилоотии донишгоҳи давлатии
технологияҳои иттилоотии Тошкент
ба номи Муҳаммад ал-Хоразмӣ

Саидзода Исроил Маҳмад -
номзади илмҳои техники, дотсент,
декани факултети механикаю математикаи
Донишгоҳи милли Тоҷикистон

Муассисаи пешбар:

Донишгоҳи технологии Тоҷикистон

Ҳимояи диссертатсия «27» феввали соли 2026, соати 14:00 дар ҷаласаи Шурои диссертатсионии 6D.KOA-049 назди Донишгоҳи техникии Тоҷикистон ба номи академик М.С. Осимӣ, баргузор мегардад. Суроға: Ҷумҳурии Тоҷикистон, 734042, ш. Душанбе, хиёбони академикҳо Раҷабовҳо, 10А. E-mail: sultonzoda.sh@mail.ru телефони котиби илмӣ: (+992) 93 700 4444.

Бо диссертатсия ва автореферати он дар китобхона ва сомонаи расмӣ Донишгоҳи техникии Тоҷикистон ба номи академик М.С. Осимӣ <http://www.ttu.tj> шинос шудан мумкин аст.

Автореферат санаи «___» _____ соли 2025 ирсол шудааст.

Котиби илмӣ
Шурои диссертатсионӣ,
номзади илмҳои техники, дотсент



Султонзода Ш.М.

МУҚАДДИМА

Мубрамии мавзуи таҳқиқот. Масъалаи татбиқи технологияҳои инноватсионӣ, яке аз масъалаҳои рушди илму маориф маҳсуб ёфта, дар риштаи илмҳои бунёдӣ, техникӣ ва эксперименталӣ дар Ҷумҳурии Тоҷикистон низ ба сафи ҳадафҳои стратегии мамлакат шомил аст.

Дар паёми охирин (28 декабри соли 2024) Асосгузори сулҳу ваҳдати миллӣ – Пешвои миллат, Президенти Ҷумҳурии Тоҷикистон муҳтарам Эмомалӣ Раҳмон оид ба зарурияти рушди информатика ва рақамикунонӣ, ворид намудани технологияи ҳозиразамонро таъкид намуданд [1].

Рушди тафаккури наврасону ҷавонон бевосита ба болоравии фарҳанги иттилоотии замони муосир боис гашта, он зина ба зина ба фазои ягонаи иттилоотӣ наздик мешавад. Дар ҷаҳони муосир тараққиёти илм ва техника бо эволютсияи воситаҳои ҳамкориҳои байни инсон ва мошин алоқаи зич дорад. Дар замони муосир интерфейси гуфтугӯӣ – ҳамкориҳои инсон ва мошин торафт маъмул мегардад. Ин аз он сабаб аст, ки барои инсон маҳз суҳан табиаттарин воситаи муошират аст. Муҳимтарин ҷузъи интерфейси суҳанронии инсон ва мошин, ки худкор мебошанд, ин шинохти овоз ва нутқ маҳсуб меёбад ва дар ин соҳа вазифаи душвортарин ба ҳисоб меравад. Инкишофи нутқ вазифаи шинохтагирии нутқи гуфтугӯӣ ва услуби нутқест, ки ба шароити набудани шакл ва мундариҷаи пешакии хабари шифохӣ ва иштироки бевоситаи суҳанварон хос аст.

Мафҳуми суҳан шакли табиӣи муоширати инсонист. Ба ҳисоби миёна, одамон дар як дақиқа 150 калима ҳарф зада метавонанд, дар ҳоле ки онҳо ба ҳисоби миёна дар як дақиқа тақрибан 50 калимаро мегӯянд, ки тахминан сеякро ташкил медиҳад. Ин нишон медиҳад, ки инсон метавонад зудтар муошират намояд ва иттилооти бештарро тавассути нутқ нисбат ба матн интиқол диҳад. Ин ҳолат одамонро водор кард, ки мошинҳоро эҷод карда, нутқи моро дақиқтар шинохта тавонанд ва иртиботи худро бо мо беҳтар созанд. Чунин мошинҳоро низомҳои гуфтугӯӣ шифохӣ меноманд ва онҳо барои муошират бо одамон пешбинӣ шудаанд.

Барои тавсифи ин мошинҳо истилоҳҳои дигар, аз қабili низомҳои фармони овозӣ ва зеҳни сунъӣ низ пешниҳод шудаанд, аммо дар масъалаи мазкур истилоҳи "низомҳои муколамаи шифохӣ"-ро истифода карда мешаванд. Гузашта аз ин, истилоҳи дигари маъмуле, ки ҳангоми паҳн шудани чунин мошинҳо пайдо шудааст, шинохтагирии овоз ва нутқи худкор буда, ба раванди табдил додани сигнали нутқ ба матни мувофиқ тааллуқ дорад.

Ширкатҳои соҳаи технологияи иттилоотӣ аҳамияти нутқро дар суҳбат дарк меkunанд ва меҳоянд сифати муоширатро бо муштариёни худ беҳтар намоянд. Бо вучуди ин, дар тӯли таърихи 30-солаи низомҳои муколамаи шифохӣ, дақиқии шинохтагирии нутқ ба қадри кофӣ баланд набуд, то ки ба сифати алтернативии инсонӣ эътимод дошта бошад. Аз ин рӯ, истифодаи низоми гуфтугӯӣ шифохӣ дар маркази таваҷҷӯҳи ширкатҳо қарор надоштаанд, то ки ба онҳо сармоягузорӣ кунанд.

Аҳамияти мавзуи таҳқиқотии мазкур, ки дар замони пешрафти технологияи инноватсионии муосир мавқеи хосаро дорад, дар сатҳи конференсияи

байналмилалӣ ва ҷумҳуриявӣ донишгоҳӣ мавриди ҷорӣ намудани низомҳои худкори шинохтагирӣ нутқ қарор дода шудааст.

Аз ин рӯ, тавассути технологияҳои инноватсионӣ барномарезӣ инъикос намудани раванди таҳия, таҳқиқ ва таҳлили низомҳои барномавии автоматикунонидашудаи шинохтагирӣ нутқ аз мубрам будани мавзӯи таҳқиқоти диссертатсионӣ шаҳодат медиҳад.

Дарачаи таҳқиқи мавзӯи илмӣ. Танҳо дар даҳсолаи охири ибтидои асри XX, баъзе ташаббусҳо вобаста ба меъёрҳои дурустии қобили қабул дар масъалаи шинохтагирӣ овоз ва нутқ ноил шуданд ва бори дигар таваҷҷӯхро бештар ба шинохтагирӣ нутқ ҷалб карданд.

Таърихи низомҳои шинохти фармони овозиро метавон ба се насл тақсим кард:

1. Дар асоси қоидаҳо ё намунаҳои аломатҳо (то охири солҳои 90-ум).
2. Дар асоси омӯзиши омӯрӣ (аз солҳои 90-ум то соли 2014).
3. Дар асоси омӯзиши амиқ (аз соли 2014).

Роҷеъ ба қори низоми иттилоотӣ ва технологияи коркарди шинохтагирӣ овоз ва нутқ, мо кӯшиш намудем, ки дар сохтани пойгоҳи додаҳо ҷӣ аз тарафи олимони ватанӣ, ҷӣ хориҷӣ ва ҷӣ муҳаққиқони соҳа ҳар маводе ё таҳқиқоти арзишманде анҷом дода шуда бошад, омӯзем, таҳлил ва таҳқиқ намоем ва аз ғояҳои онҳо истифода намоем.

Сатҳи азхудшудаи мавзӯи диссертатсионӣ мазкур асосан ба таҳқиқоти олимони хориҷи наздик ва дур алоқаманд буда, дастовардҳои илмӣ онҳо дар ҷодаи коркарди системаҳои иттилоотӣ асоси методологияи таҳқиқотро ташкил додаанд, аз ҷумла, таҳқиқоти Дмитрий Богданов, Кенхен Чо, Йошуа Бенгио, Ян Хоровский, Дмитрий Сердюк, Филимон Бракељ, Жанет М. Бейкер, Ли Дэн, Джеймс Гласс, Санджив Худанпур, Уильям Чан, Навдип Джейтли, Чжифэн Чен, Анджули Каннан, Ярин Гал, Зубин Гахрамани, Феликс Герс, Юрген Шмидхубер, Фред Камминс, Виталий Липчинский, Раджат Райна, Ананд Мадхаван, Эндрю Янг, Карл Уилмер Шольц, Джеймс С. Ирвин, Юнхуэй Ву, Майк Шустер, Чжифэн Чен, Alex Graves, Santiago Fernández, Faustino Gomez, Jürgen Schmidhuber ва дигарон.

Коркарди амсилаҳо, усулҳо ва алгоритмҳо барои шинохти воҳидҳои овозӣ ва нутқи сарбаст дар асоси шабакаҳои нейронӣ ва омӯзишимошинӣ таърихи на он қадар зиёд дорад. Дар осори илмӣ як зумра олимони ба монанди Gaikwad S.K., Gawali B.W., Yannawar P.A., Abdel-Hamid O., Johnson E.M., Богустов А.А., Парамонов П.А., Усманов З.Д., Исмоилов М.А. ва Худойбердиев Х.А. воситаҳои асосии зеҳни сунъӣ, сохтори корпуси овоз, технологияи коркарди амсилаҳои амалии омӯзиши мошинӣ, ки минбаъд дар низомҳои худкори шинохти нутқ ба забони табиӣ дар асоси шабакаҳои нейронӣ пешниҳод шудаанд.

Дар осори олимони рус Барский А. Б., Богустов А. А., Меденников И.П., Гребнов С. В., Тампель И. Б., Карпов А. А. ва дигарон масъалаи коркарди худкори воҳидҳои овозӣ ва шинохтагирӣ нутқ натиҷаҳои дақиқ ва сахт ба даст оварда шудааст. Натиҷаҳои таҳқиқоти олимони номбаршуда имкониятҳои усулҳои барномарезии динамикӣ, таҳлили нутқи сарбаста пешниҳод гардиданд. Илова бар ин, масъалаи шинохтагирӣ фармонҳои овозӣ аз қорбар новобаста низ ба роҳ монда шудаанд.

Масалан, дар рисолаи диссертатсионӣ худ Барский А.Б., қайд менамояд, ки “ба ҳар ҳол пас аз пайдоиши омӯзиши амиқ дар вазифаҳои шинохтагирӣ нутқ,

ширкатҳо низ ба истифодаи чунин система дар соҳаҳои тиҷоратии худ манфиатдор мебошанд. Тавассути истифодаи омӯзиши амиқ, шинохтагирии нутқ ба қадри кофӣ мукамал гардид, ки дар ҳолатҳои истифодаи воқеӣ ва тиҷоратӣ муфид ҳисобида мешавад” [3].

Дар мавриди таҳқиқоти олимони ватанӣ ва дастовардҳои илмӣ онҳо дар ҷодаи коркарди системаҳои иттилоотӣ, ки асоси методологияи таҳқиқотро ташкил додаанд, дар мисоли дастпарварони мактаби илмӣ олимони шинохтаи риштаи технологияи информатсионӣ академики АМИТ – Усмонов З.Ҷ., кормандони пешбари илмӣ Донишгоҳи технологияи Тоҷикистон – Исмоилов М.А., Зарипов С.А., Исмоилова Р.М., Ғуломсафдаров А.Ғ., кормандони илмӣ Донишгоҳи техникии Тоҷикистон ба номи академик М.С. Осимӣ – Солиев О.М., Худойбердиев Х.А., Довудов Г.М., Қосимов А.А., Ашурзода Б.Х. ва дигарон дида метавонем [4].

Яке аз натиҷаҳои назаррас дар самти шинохти калимаҳои калидӣ дар нутқ бо забони тоҷикӣ аз тарафи Ашурзода Б.Х. бо роҳбарии Худойбердиев Х.А. дар таҳияи амсилаи низоми шинохти нутқ бо интиҳоби калимаҳои калидӣ аз нутқи пайваст бо забони тоҷикӣ ба ҳисоб меравад. Дар раванди таҳқиқот низоми шинохти калимаҳои калидӣ дар асоси амсилаҳои ноаёнӣ Марков муайян карда шудааст, инчунин дар ин кори таҳқиқоти аз ҷумла калимаҳои додашударо, ки дар анбори маълумотҳо мебошад дарёфт менамояд.

Бояд қайд кард, ки усулҳо ва афзорҳои дар рисолаи мазкур зикргашта, дар асоси назарияи муосири низоми объектҳои динамикӣ бунёд гаштаанд. Дар рафти иҷрои таҳқиқоти диссертатсионӣ, муаллиф ба принсипҳои асосии таҳияи низомҳои идоракунии амсиласозии компютерӣ таъя намуда, ҳалли алгоритми масъалаи шинохтагирии нутқро ёфтааст.

Таҳқиқоти мазкур бо он ҷолиб аст, ки дар асоси омӯзиши ҳамаҷонибаи платформаҳои веб-замимаҳо ва системаҳои мавҷудаи идоракунии барномаҳои компютерӣ таснифоти мукаммали онҳо рӯйи кор омадааст. Чунон ки вақтҳои охир шинохтагирии нутқ ҳамчун қадами аввалин дар низоми иртиботи инсон бо компютер бо қобилият ва арзиши ҷолиби тиҷоратӣ таваҷҷӯҳи бештарро ба худ ҷалб кардааст.

Бо афзоиши эҳтиёҷот ва беҳбудии иртибот бо муштариён, бисёре аз ширкатҳо бо назардошти сатҳи талаботи амалишавии чунин системаҳо дар кишварҳо, ба монанди ИМА, Россия, Чин, ва ғ. ҳадди ақал бо низоми фармонҳои овозӣ байни худ озмоишҳо барпо намудаанд. Мувофиқи Gartner (ширкати таҳқиқоти дар соҳаи технологияи иттилоотӣ), то соли 2021 15% -и тамоми муоширати муштариён пурра аз ҷониби зеҳни сунъӣ идора карда мешуд.

Вобаста ба масъалаи мазкур таҳқиқоти дигар низ вучуд дорад ва он нишон медиҳад, ки 87% иштирокчиёни самти тиҷоратӣ боварӣ доранд, ки дар асоси зеҳни сунъӣ гуфтугӯ ба сатҳи баланди қаноатмандӣ дар риштаи тиҷоратии онҳо оварда мерасонад.

Бо вучуди тамоми кӯшишҳо ва беҳбудӣҳо дар самти шинохтагирии нутқ, дар ҳаҷми зарурӣ ба даст овардани он ҳанӯз душвориҳо дорад, зеро дар ин самт якчанд омилҳои таъсиррасон ҷой доранд.

Чунин омилҳо, шиноختан ва пурра дарк кардани нутқи инсон бо воситаи техникаи ҳисоббарорро душвор мегардонанд. Масалан, аз сабаби тағйирпазирии

сигнали нутқ, таҳлили нутқи инсон чун сухани давомдор ба ҳисоб меравад, барои мошинҳо кори душвор аст.

Бо вучуди беҳбудии назаррас дар шинохтагии овоз ва нутқи тоҷикӣ, беҳбудии барои забонҳои дигар ҳанӯз заруранд. Муваффақияти шинохтагии нутқ дар як забон, ба забонҳои дигар паҳн намешавад.

Аз ин рӯ, дар ин таҳқиқот кӯшиши коркарди низомии шинохтагии овоз ва нутқро истифода шуда, дар яке аз самтҳои тижорати ширкати молиявӣ ва ё дар раванди таълими муассисаҳои таҳсилоти олии касбӣ амалӣ карда мешавад.

Тибқи натиҷаҳои таҳқиқотӣ муайян гашт, ки шумораи зиёди корбарони компютер аз низомҳои пешрафтаи шинохти нутқ бо забони табиӣ лоиҳаҳои “Langsoft”, “Lexical FreeNet”, “LingSoft”, “MonoConc/ParaConc”, “Mystem”, “netXtract”, “Ngram Statistics Package – NSP”, “Open Office Org”, “Sakrament Text-to-Speech Engine”, “Speech Synthesis and Recognition Laboratory”, “Speech technology”, “ASRaas”, “TTSaaS SaluteSpeech API”, “Sputnik Voice”, “Audiogram”, “SmartNLP”-ро истифода мебаранд. Шумораи зиёди лоиҳаҳо дар шакли бархат (онлайн) ва дар ҳолати имконпазир бидуни тамос (офлайн) истифода карда мешаванд. Низомҳои номбаршуда дар асоси шабакаи нейронӣ ва усулҳои омӯзиши мошинӣ коркард ва ба истифода дода шудаанд. Лоиҳаҳои ба монанди “SteosVoice”, “Lovo.ai”, “Zvukogram”, “FreeTTS”, “Robivox”, “PlayHT”, “Deepgram”, “Murf.ai” дар бозори низомҳои шинохтагии нутқ барои зиёда аз 40 забонҳои табиӣ васеъ истифода карда мешаванд. Аммо лоиҳаҳои номбар шуда барои шинохти нутқи тоҷикӣ дастгирӣ карда наметавонанд. Дар асоси гуфтаҳои боло мубрами коркарди лоиҳаи шинохтагии фармонҳои овозӣ ва нутқ бо забони тоҷикӣ низ зарурати истифодаи шабакаҳои нейронӣ ва усулҳои омӯзиши мошинӣ пеш назар мерасад.

Пеш аз ҳама, яке аз афзалиятҳои асосии истифодаи шинохтагии нутқ кам кардани захираҳои инсонӣ дар хизматрасонӣ ба муштариён аст, ки дар натиҷаи ин хароҷоти молиявӣ нисбати кормандони соҳа кам мешавад. Илова бар ин, ёвари овозии фардӣ метавонад садоқат ва қаноатмандии муштариёро баланд бардорад.

Таҳқиқоти солҳои охир нишон медиҳанд, ки омӯзиши амиқ дар ширкати ASR муваффақшуда маҳсуб меёбад. Дар байни сохторҳои шабакавии мавҷудбуда, тавачҷӯҳи ин таҳқиқот ба CNNs, RNNs бештар мебошад ва ҷузъҳои муҳимро дар системаҳои омӯзиши амиқ мавриди муҳокима қарор медиҳанд. Ин ҷузъҳо сохтори амсилавӣ, маҷмӯи таълимии калонҳаҷм, миқёси ҳисоббарорӣ, коркарди пешакии додашудаҳо ва танзими гиперпараметрҳо мебошанд. Инчунин, таъсири ин ҷузъҳо ба дурустии шинохти нутқ зикр карда мешавад. Аз ҷумла, ин кор тавассути омӯзиши таҷрибавии дақиқи функционалии шабакаҳои нейронӣ дар вазифаҳои шинохти нутқ дастгирӣ карда мешавад. Дар навбати худ, боз ду амсилаи гуногуни шабакаи нейронӣ барои пешгӯии транскрипсияи нутқ аз аудiotачрибаҳо гузаронида шудаанд.

Аз нуктаи назари сатҳи баланд, ҳадафи ниҳии ин таҳқиқот кушодани роҳ ба рушди амсилаи миқёспазир ва амалии шинохти овоз ва нутқ, ки маҳдудиятҳои воқеиро ба инобат мегирад, яъне амсилае, ки ҳам корбарон ва ҳам ширкатҳо метавонанд аз он баҳра баранд.

Солҳои охир дар низомии иттилоотӣ хизматрасонии коммуникатсионӣ қисми зиёди муҳаққиқон барои коркарди иттилоот аз усулҳои гуногун ва амсилаҳои

математикии “вақттаксимкунӣ”-и ба назарияи хизматрасонии оммавӣ асосноккардашуда истифода бурдаанд.

Амсилаи математикӣ-компютерии низоми иттилоотӣ хизматрасонии пойгоҳи шинохтагирии нутқи тоҷикӣ низ дар асоси қонунҳои назарияи хизматрасонии оммавӣ, тарзи кори шабакаҳои телекоммуникатсионӣ ва принципҳои асосии амсиласозии математикӣ таҳия гашта, ғояи марказии ба муҳити шабакавӣ мутобиқгардонии он аз таҳқиқоти “ЗИАО-ШН”-и олимони давлатҳои аврупо иқтибос гардидааст.

Робитаи таҳқиқот бо барномаҳо (лоиҳаҳо) ва мавзӯҳои илмӣ. Натиҷаҳои назариявӣ ва амалии дар таҳқиқот ба даст оварда, ки дар муайян намудани алгоритми дақиқии шинохтагирии фармонҳо дар нутқ бо забони тоҷикӣ талаффуз шуда ва татбиқи он дар низоми босуръати кофӣ барои таҷрибавӣ нигаронидашуда, дар доираи амалисозии “Бистсолаи омӯзиш ва рушди фанҳои табиатшиносӣ, дақиқ ва риёзӣ дар соҳаи илму маориф” барои солҳои 2020-2040 ва Барномаи “Ҳукумати электронӣ” тарҳрезӣ гаштаанд. Илова бар ин, натиҷаҳои рисолаи диссертатсионӣ барои иҷрои бандҳои “Барномаи давлатии амалӣ намудани технологияҳои иттилоотӣ-коммуникатсионӣ дар муассисаҳои таҳсилоти умумии Ҷумҳурии Тоҷикистон” барои солҳои 2018-2022, Консепсияи ташаккули Ҳукумати электронӣ дар Ҷумҳурии Тоҷикистон, Қонуни Ҷумҳурии Тоҷикистон “Дар бораи забони Давлатии Ҷумҳурии Тоҷикистон”, Қонуни Ҷумҳурии Тоҷикистон “Дар бораи иттилоот” ва пешбарии супоришҳо дар нақша-чорабиниҳои дар Стратегияи давлатии “Технологияҳои иттилоотӣ-коммуникатсионӣ барои рушди Ҷумҳурии Тоҷикистон”, Стратегияи рушди инноватсионии Ҷумҳурии Тоҷикистон барои давраи то соли 2020, Стратегияи миллии рушди Ҷумҳурии Тоҷикистон барои давраи то соли 2030 мувофиқат мекунад, ки мубрамии мавзӯи илмиро тасдиқ менамоянд [2].

ТАВСИФИ УМУМИИ ТАҲҚИҚОТ

Мақсади таҳқиқот – таҳияи амсилаҳои математикӣ ва компютерӣ, усулҳои коркарди нутқ дар асоси шабакаҳои нейронӣ ва алгоритмҳое, ки барои шинохти худкори нутқи тоҷикӣ истифода карда шуда, ба роҳ мондани муколама бо забони табиӣ дар низомҳои “инсон-мошин” ва маҷмӯи барномаҳои компютерӣ.

Вазифаҳои таҳқиқот:

-омӯзиш, таҳқиқ ва таҳлили интишороти мавҷуда оид ба технологияҳои сохтани низоми иттилоотӣ ва пойгоҳи додаҳо барои шинохтагирии нутқи тоҷикӣ;

-мураттабсозии мақсад ва вазифаҳо, таҳлили ҳолати таҳқиқот ва адабиёти мавҷуда, таҳияи қарорҳои алгоритмӣ ва нармафзор, арзёбии таҷрибавӣ ва самаранокии қарорҳои таҳияшудаи таҳқиқот, санҷиш ва таҳлили натиҷаҳо;

-ба сифати усулҳои таҳқиқот, истифодаи усулҳои коркарди сигналҳои рақамӣ, назарияи эҳтимолият ва амсилаи математикӣ, омӯзиши мошинӣ, лингвистикаи амалӣ ва коркарди таъминоти барномавӣ;

- коркарди амсилаҳои математикӣ ва компютерии дастраси умум ба пойгоҳи додашудаҳои архитектураи шинохтагирии овоз ва нутқи тоҷикӣ;

-коркарди алгоритми пайвандсозии гипер матнҳои маҳаллӣ ба сохтори барномавию объектгароии амсилаи шинохтагирии нутқ;

-баҳодиҳии сифати кори системаи таҳияшудаи эътирофи шинохти овоз ва нутқи тоҷикӣ, дар асоси хусусиятҳои гуногуни корбарони низоми иттилоотӣ.

Объекти таҳқиқот – низоми автоматикунонии иттилоотӣ ва замимаҳои пойгоҳи додаҳои шинохтагирии нутқи тоҷикӣ дар намуди архитектураи “муштарӣ-сервер-муштарӣ”.

Мавзӯи таҳқиқот – усулҳои самараноки коркард, таҳқиқ, таҳлил ва татбиқи системаи барномавӣ-алгоритмӣ ва воситаҳои таъминоти барномавӣ бо эътирофи автоматикунонии шинохтагирии овоз ва нутқи тоҷикӣ.

Асосҳои назариявии таҳқиқот. Дар доираи ҳалли масъалаҳои муҳрам оид ба технологияҳои коркарди нутқ, алалхусус шинохти худкори нутқ бо забони тоҷикӣ методологияи таҳлили омӯрӣ ва амсилаи пинҳони Марков мавриди истифода қарор дода шудаанд, ки дараҷаи онҳо на он қадар босамар мебошанд. Инчунин дар доираи таҳқиқот корпусҳои нутқи забонҳои русӣ ва форсӣ бо мақсади шинохти худкори нутқи забони тоҷикӣ аз ташхис гузаронида шуданд, ки дараҷа ва эътимоди аз талабот хеле пастро пайдо карданд. Дар таҳқиқот методологияи зехни сунъӣ, яъне усулҳои муосири омӯзиши мошинӣ ва шабакаҳои нейронӣ барои босифат ва босуръат шинохтагирии фармонҳо дар нутқи бо забони тоҷикӣ талаффузшуда мавриди истифода қарор дода шуданд. Бо мақсади ба даст овардани дараҷаи баланди шинохти нутқ ва эътимоднокии он бори аввал амсилаи шабакаи нейронӣ барои омӯзиши мошинӣ бо воҳидҳои овозии бо забони тоҷикӣ талаффуз ва сабтшуда, коркард карда шудааст. Усулҳои омори математикӣ ва назарияи эҳтимолият барои баҳодиҳии самаранокии раванди шинохтагирии фармонҳо дар нутқи бо забони тоҷикӣ талаффуз шуда, истифода шудаанд.

Навгониҳои илмӣ. Дар натиҷаи кори таҳқиқотӣ чунин навгониҳои илмӣ ба даст оварда шуданд:

- бори нахуст дар заминаи пойгоҳҳои додаҳои мавҷуда бо баҳисобгирии хусусиятҳои маҳаллӣ барои шинохтагирии нутқи тоҷикӣ дар шакли як низоми мукаммали таҳқиқотӣ дар асоси амсилаҳои математикӣ ва компютери коркардшуда таҳия гардиданд;

- барои сохтани хусусиятҳои иттилоотӣ, ки аз шабакаи амиқи нейронӣ бо гулӯи танг гирифта шудаанд, бо истифодаи мутобиқшавӣ ба гуфтор ва шароити акустикӣ фарқ мекунад, мавриди коркард қарор дода шудаанд;

- барои баланд бардоштани сифати шинохтагирии нутқ алгоритми ду марҳилаи ибтидоии акустикаи таълимӣ таҳия шуда, дар асоси шабакаҳои амиқи нейронӣ бо назардошти симои рақамии фармонҳои овозӣ дар маҷмӯи омӯзиш қарор дода, пешниҳод карда шуданд;

- сохтори мантиқии барномавӣ, равандҳои корӣ ва технологияҳои дастрас барои шинохтагирии фармонҳои овозӣ бо забони тоҷикӣ коркард ва татбиқи онҳо дар модели компютерӣ қарор дода шудаанд;

- бори аввал барои ҳалли масъалаи шинохти нутқи тоҷикӣ дар асоси шабакаҳои нейронӣ дар маҷмӯи барномаҳои компютери Tajik-AVCR бо пешниҳоди амсилаи математикӣ равандҳои корӣ ва сохтори мантиқии низоми шинохтагирии нутқ дар асоси CASE-воситаҳо таҳия карда шуданд.

Нуктаҳои ба ҳимоя пешниҳодшаванда:

1. Амсилаҳои математикӣ ва компютерӣ, усули сохтани хусусиятҳои иттилоотии нутқ ва симои рақамии он дар асоси шабакаи амиқи нейронӣ муайян

карда шуданд.

2. Усулҳои сохторбандии низомии шинохтагирии нутқи забони тоҷикӣ ва омӯзиши мошинӣ дар амсилаи TajNN-SAT ташкил ва пешниҳод карда мешавад.

3. Комплекси барномаҳо барои шинохтагирии фармонҳо бо забони тоҷикӣ коркард карда шуд, ки дар равандҳои кори соҳаҳои гуногуни ҷамъият ба монанди таълим, тиб ва сайёҳӣ босамар мавриди истифода қарор доранд.

Аҳамияти назариявӣ ва амалии таҳқиқот. Аҳамияти назариявии рисолаи мазкур такмил додани алгоритмҳои мавҷуда ва таҳияи алгоритмҳои нави омӯзиши амсилаҳои акустикӣ дар асоси шабакаҳои амиқи нейронӣ барои мушкилоти шинохтагирии овоз ва нутқ, инчунин таҳия ва ба таври таҷрибавӣ омӯختани усули нави архитектураи хусусиятҳои иттилоотӣ, ки аз алгоритмҳои истифодашуда бартарӣ дорад, иборат аст. Аҳамияти амалии таҳқиқот дар истифодаи асбобҳои алгоритмӣ ва нармафзори таҳияшуда, ҳангоми сохтани системаи шинохтагирии овоз ва нутқи тоҷикӣ, нишон додани сифат ва суръати ба кадри кифоя баланд, барои истифода дар чунин вазифаҳои амалӣ, дар мисоли нақлунии автоматии архивҳои фонограмма, ҷустуҷӯи калимаҳои калидӣ дар ҷараёни нутқ, гурӯҳбандии сабтҳо аз рӯи мавзӯро таъмин менамояд. Натиҷаҳои асосии дар рисолаи номзадӣ бадастомада амалӣ карда мешаванд: дар барномаи "Системаи шинохтагирии овози забони тоҷикӣ", дар раванди таълими ДДХБСТ, барномаи DEM.tj ҳангоми иҷрои корҳои таълимию методӣ аз тарафи ҳайати профессорону омӯзгорон ва донишҷӯён ва дар раванди таълими Донишкадаи саноат ва хизматрасонӣ.

Дарачаи эътимоднокии натиҷаҳои диссертатсия. Натиҷаҳои назариявӣ ва амалии таҳқиқот мувофиқ мақсад ва вазифаҳои он ба методологияи зеҳни сунъӣ асос ёфтаанд. Дар ин раванд усулҳои муосири таҳлили системавӣ, концепсияи илмӣ таҳлили компютерӣ ва амсиласозии компютерӣ барои омӯзиши мошинии шабакаҳои нейронӣ бо таҷрибаи шахсии муаллиф асоснок буда, пояи бозэтимод ва мазмуну мундариҷаи диссертатсия, ҳулосаҳо ва тавсияҳои дар он ифодагардидаро ташкил додаанд. Натиҷаи корҳои илмӣ-таҳқиқоти дар озмуни ҷумҳуриявии "Илм – фуруғи маърифат" соли 2024-2025, дар номинатсияи "Зеҳни сунъӣ ва барномасозӣ" миёни ходимони илмӣ, аспирантону докторантон, омӯзгорони муассисаҳои таълимӣ, муҳандисон, конструкторон ва намоёндагони касбу кори гуногун муаррифи гардида, сазовори ҷои дуюм гардид. Аз рӯи натиҷаҳои бадастомада комплекси таъминоти барномавии Tajik-AVCR коркард, такмил ва татбиқ карда шуданд. Натиҷаҳои асосии таҳқиқот аз рӯи соҳаи илм ва маориф дар Донишгоҳи давлатии ҳуқуқ, бизнес ва сиёсати Тоҷикистон ва Донишкадаи саноат ва хизматрасонӣ, аз рӯи соҳаи тиб дар ҚДММ "Баҳри Шифо" ва ҚДММ "Ал-Финур", аз рӯи соҳаи сайёҳӣ дар ҚДММ "Марвел Вита Травел" ва ширкати сайёҳии "Панҷистон" дар мавриди истифода қарор дода шудаанд.

Мутобиқати диссертатсия ба шиносномаи ихтисоси илмӣ. Таҳқиқот дар доираи ихтисоси 6D070400 – Техникаи ҳисоббарор ва таъминоти барномавӣ (6D070402-Мошинҳои ҳисоббарор, комплексҳо ва шабакаҳои компютерӣ):) ба иҷро расонида шудааст. Диссертатсия аз рӯи бандҳои зерини Шаҳодатномаи номгуӣ ихтисосҳои коргарони илмӣ анҷом дода шудааст:

банди 2 – *"Таҳлили назариявӣ ва таҷрибавии кори мошинҳои ҳисоббарор, комплексҳо ва шабакаҳои компютерӣ бо мақсади беҳтар намудани хусусиятҳои техникӣ-иқтисодӣ ва амалиётҳои онҳо"* амсилаи компютерӣ ва комплекси

барномаҳои компютери коркардшуда, мувофиқи банди мазкур мебошад, ки он дар асоси хусусиятҳои воситаҳои зеҳни сунӣ ва шабакаҳои нейронӣ доранд ва барои баҳодиҳии боэътимоди усулҳо ва кори комплексҳоро аз санҷиш гузаронад;

банди 3 – *“Таҳияи усулҳо ва алгоритмҳои илмӣ оид ба ташики коркарди арифметикӣ, мантиқӣ, рамзӣ ва махсуси додаҳо, нигоҳдорӣ ва воридоту содироти иттилоот”* амсилаи компютери коркардшуда мувофиқи банди мазкур мебошад, ки он барои ташики коркарди самараноки додаҳои овозӣ имконият фароҳам меорад ва бо мақсади шинохтагрии худкори фармонҳо дар воҳидҳои нутқи бо забони тоҷикӣ талаффузшуда, истифода шавад;

банди 4 – *“Коркарди усулҳо ва алгоритмҳои илмӣ оид ба ташики коркарди додаҳои параллелӣ ва тақсимишуда, системаҳои ҳисоббарори бисёрпротсессорӣ, бисёрмошинӣ ва махсус”* комплекси таъминоти барномавии коркардшуда, ки дар асоси усулҳо ва алгоритмҳои коркарди додаҳои намудашон махсуси рақамӣ барои шинохтагрии худкори фармонҳо дар нутқи забони тоҷикӣ самаранок истифода шудаанд, шомили бди мазкур мебошад;

Саҳми шахсии довталаби дараҷаи илмӣ дар мустақилона навиштани диссертатсия, иштироки бевосита дар масъалагузориҳои мавзӯ, таҳқиқоти назариявӣ, алгоритмсозӣ, барномарезӣ, амсиласозӣ, баргузориҳои озмоишҳои компютерӣ, таҳлил ва ҳулосабарориҳои ҷамъбасти ифода ёфтааст. Усулҳо ва алгоритмҳои шинохти овоз ва нутқро коркард намуда, баҳодиҳии таҷрибавии самараноки усулу алгоритмҳои таҳияшуда гузаронида шудааст. Воситаҳои нармафзор таҳия карда шудаанд, ки як қисми асосии системаи шинохтагрии овоз ва нутқи тоҷикӣ мебошанд. Нуқтаҳои асосии илмӣ ва ҳулосаҳои диссертатсионӣ тавассути усули амсиласозии компютерӣ ва модулҳои барномавии дар шакли комплекси барномаҳои объектгаро таҳияшуда асоснок гардида, тасдиқи ҳудро ёфтаанд.

Тасвиб ва амалисозӣ. Натиҷаҳои гузориши рисолаи номзадӣ дар конференсияҳои илмию методии зерин муҳокима, таҳия ва дар маҷмӯи мақолаҳои конферонсҳо интишор гардидаанд: конференсияи илмӣ-амалии олимону профессорон ва муҳаққиқони ҷавони ДДҲБСТ (солҳои 2019-2020-2021-2022), конференсияи илмӣ-назариявии “Илмҳои бунёдӣ ва амалӣ: ҳолат ва рушди соҳа” (Петрозаводск, 2022), конференсияи илмӣ-назариявии «Нақши илм дар ташаккули давлатдорӣ миллӣ» (Хучанд, 2018); «Рушди фанҳои табиатшиносӣ, дақиқ ва риёзӣ дар низомии ташаккулёбии иқтисодиёти рақамӣ» (Хучанд, 2020); конференсияи илмӣ-амалии «Илм ва инноватсия дар низомии татбиқи ҳадафҳои стратегияи миллӣ» (Хучанд, 2021); конференсияи илмӣ-методии «Усулҳои инноватсионии таълим: ҳолати муосир ва дурнамои тақомул» (Хучанд, 2021) ва «Консепсияи миллии тарбия – асоси тарбияи насли созандаю бунёдкор» (Хучанд, 2021); конференсияи илмӣ-методии «Муаммоҳои мубрами методикаи таълими фанҳои риёзӣ, дақиқ ва табиӣ дар МТОК: назария ва амалия»; (Хучанд, 2022); конференсияи илмӣ-амалии «Ҳамгироии илм ва истеҳсолот баҳри татбиқи ҳадафҳои стратегияи миллӣ» (Хучанд, 2022); конференсияи илмию методии «Ваҳдати миллӣ – омили муҳими рушди давлат ва ҷомеа» (Хучанд, 2022) .

Интишорот аз рӯйи мавзӯи диссертатсия. Натиҷаҳои таҳқиқот дар 18 интишороти муаллиф инъикос ёфтаанд, ки аз онҳо 6-тояшро мақолаҳои илмӣ дар маҷалаҳои тақризшавандаи Комиссияи олии аттестиатсионии назди Президенти

Ҷумҳурии Тоҷикистон банаҳррасида, боқимонда мақолаҳои дар дигар нашрияҳо ба нашр расида, маводи конференсияҳои илмӣ ташкил додаанд. Аз он ҷумла, Шаҳодатнома оид ба бақайдгирии асарҳои илмӣ ва ҳуқуқи муаллифӣ барои лоиҳаи Web-замимаи “Системаи шинохтагии фармонҳои тоҷикӣ” аз 23.02.2024, №188 аз тарафи Вазорати фарҳанги Ҷумҳурии Тоҷикистон ва Шаҳодатнома оид ба бақайдгирии давлатии захираҳои иттилоотӣ барои лоиҳаи Web-замимаи “Системаи шинохтагии овоз бо забони тоҷикӣ” аз 10.07.2024, №4202400519 аз тарафи “Маркази миллии патенту иттилоот”-и назди Вазорати рушди иқтисод ва савдои Ҷумҳурии Тоҷикистон гирифта шудаанд.

Соҳтор ва ҳаҷми диссертатсия. Диссертатсия аз қисмҳои “Муқаддима, “Тавсифи умумии кор”, се боб, “Хулоса” ва “Тавсияҳо”, “Рӯйхати адабиёти истифодашуда” ва “Рӯйхати интишороти илмӣ довталаби дараҷаи илмӣ” ва се замима иборат мебошад. Ҳаҷми умумии диссертатсия аз 165 саҳифаи матн компютери бо ёрии протсессори матнии Microsoft Word ҳуруфчиншуда иборат буда, 31 расм, 2 диаграмма ва 16 ҷадвалро дар бар гирифтааст. Рӯйхати адабиёт фарогири 154 номгуй мебошад.

МУНДАРИҶАИ АСОСИИ КОР

Дар муқаддима мубрамияти мавзӯи интиҳобшуда таъкид шуда, ҳадаф ва вазифаҳои асосии он асоснок карда шудаанд, аҳамияти илмӣ ва амалии таҳқиқот нишон дода шудааст. Мундариҷаи мухтасари рисола баён карда шудааст.

Дар **боби 1** “*Асосҳои назариявии амсилаҳо, усулҳо ва алгоритмҳои шинохти нутқ*” масъалаҳое, ки бо муайян кардани хусусиятҳои асосии тарҳрезӣ ва мушкилоти ҳалталаби татбиқи низомҳои коркарди иттилоот ба забони тоҷикӣ алоқаманданд, баррасӣ карда мешаванд.

Дар **§ 1.1** маълумоти умумӣ оид ба *таснифоти низомҳои шинохти нутқ* натиҷаҳои илмӣ-амалии истифодаи низомҳои шинохти нутқ тасниф карда шудаанд. Таҳқиқот нишон медиҳад, ки шинохти нутқ аввалин қобилиятест ва тавассути он низоми зеҳни сунъӣ бояд бо инсон гуфтугӯи инсонӣ дошта бошад. Шинохти нутқ метавонад овози инсонро эътироф кунад ва нутқро ба матн табдил диҳад, ки ин як шакли маълумоти бо мошини ҳисоббарор хондашаванда маҳсуб меёбад. Ин таҳқиқот идомаи корҳои қаблӣ дар омӯзиши амиқ барои шинохти нутқ мебошад.

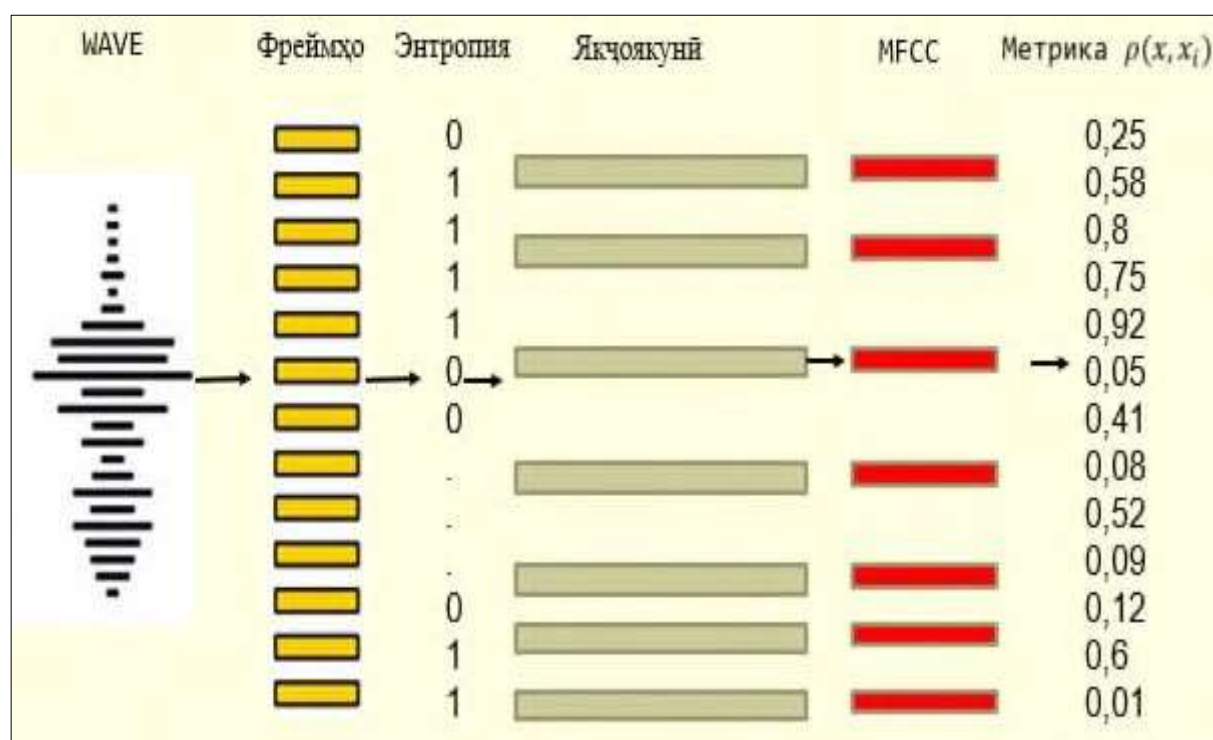
Таҷрибаи солҳои охир нишон медиҳад, ки технологияи шинохти нутқ бо истифодаи технологияҳои рақамӣ дар байни ширкатҳои мобилӣ босуръат азхуд гардида, татбиқи он дар соҳаҳои гуногун, дар мисоли муассисаҳои таълимӣ, бонкдорӣ, маркетинг, тандурустӣ, сайёҳӣ возеҳ ба чашм мерасад.

Дар **§ 1.2** *таснифоти амсилаҳо ва усулҳои шинохти нутқ* дар раванди шинохтан тавассути муайян кардани хусусиятҳои шабеҳ тавассути тақсим кардани тасвирҳо ба синфҳои умумӣ оварда шудаанд. Ҷиҳати асосӣ дар ин усул дар он аст, ки тасвирҳо аз як қатор хосиятҳо ва ё хусусияти умумӣ иборат буда, ба як синф мансуб мебошанд. Ин хосиятҳои муштаракро ба хотираи низом метавон ворид намуд.

Барои ба амал баровардани амсилаи татбиқшавандаи коркарди сигналҳо ва табдили онҳо ба овоз, алгоритми соҳтор чунин намудро мегирад. Чунон ки аз расми

1 дида мешавад, алгоритмҳо барои коркарди сигнали нутқ дар амсилаи дарккунӣ, бояд ҳамон низоми концепсияҳо ва муносибатро, ки шахс истифода мебарад, истифода намояд. Амсилаи акустикӣ имкон медиҳад, ки сегменти баромадро аз ҷиҳати шабоҳат дар сатҳи садо ба назар гирифта, симои рақамии нутқ ба даст оварда шавад. Барои ҳар як садо, аз ибтидо амсилаи мураккаби омӯрӣ сохта мешавад, то ки он натиҷаи ин садоро дар мавриди сабти он тасвир кунад.

Низоми коркардшаванда барои шинохти фармонҳои овозии дар алоҳидагӣ талаффузгашта пешбинӣ мегардад. Шинохти нутқи равон аз доираи таҳқиқоти диссертатсия берун аст ва усулҳои мукаммали коркардро талаб мекунад. Низоми шинохти коркардшаванда аз рӯи таъинот низоми фармониро ифода менамояд. Алгоритми шинохтро усули барномасозии динамикӣ ташкил медиҳад. Аз рӯи намуди воҳидҳои сохторӣ низом бояд калимаҳои ҷудогоноро шинохта гирад. Чунин талаботи дар боло қайдшуда, гузориши масъалаи таҳқиқотиро ташкил медиҳанд.

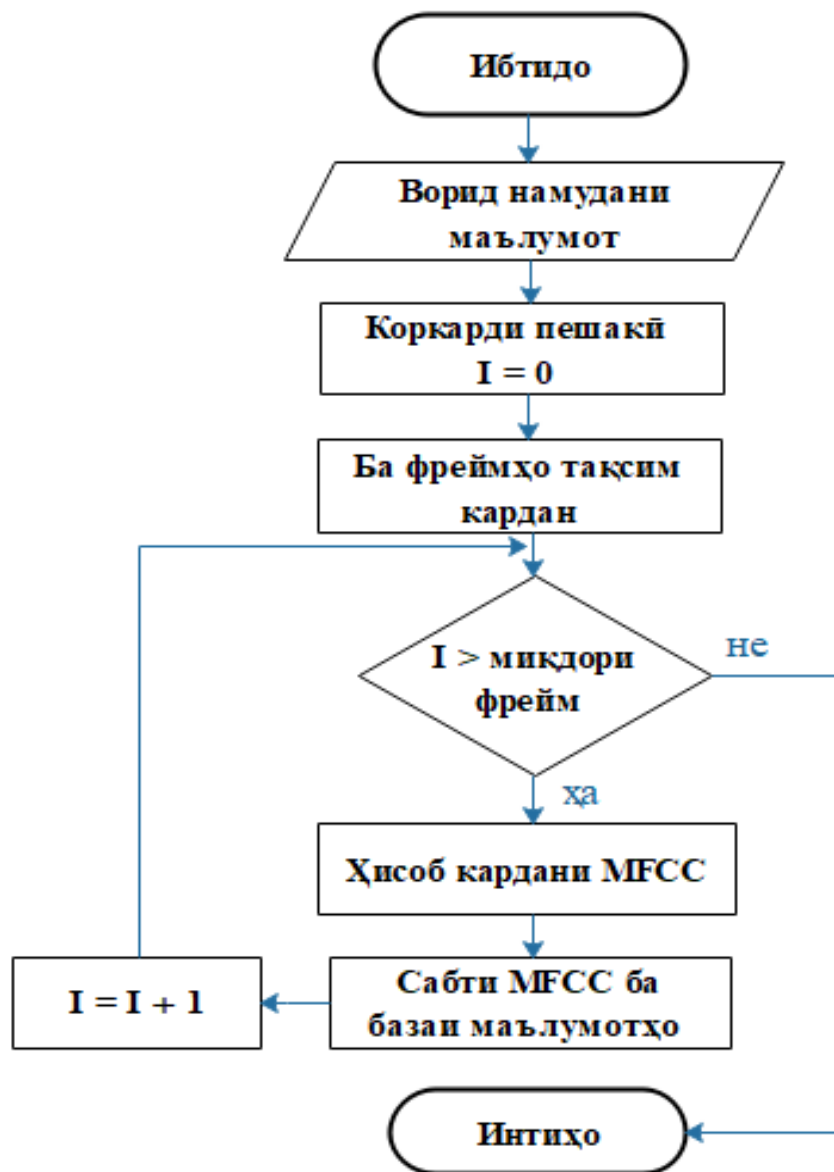


Расми 1 - Алгоритми асосии шинохти рақамии нутқ

Дар § 1.3 истифодаи амалии алгоритми шинохти нутқ дар асоси тағйирёбии динамикии вақт дар мисоли нутқи забони тоҷикӣ натиҷаҳо пешниҳод карда шудаанд. Масъалаи коркарди воҳидҳои овозӣ ва талаффузи худкори нутқро дар асоси мувофиқати намуна дарбар мегирад. Хотираи низоми шинохт тасвирҳои зиёдеро, ки ба як синф тааллуқ доранд, нигоҳ медорад. Вақте ки дар низом тасвирҳои нав пайдо мешаванд, низом онҳоро як ба як бо тасвирҳои, ки аллакай дар хотираи он нигоҳ дошта мешаванд, муқоиса мекунад (расми 2).

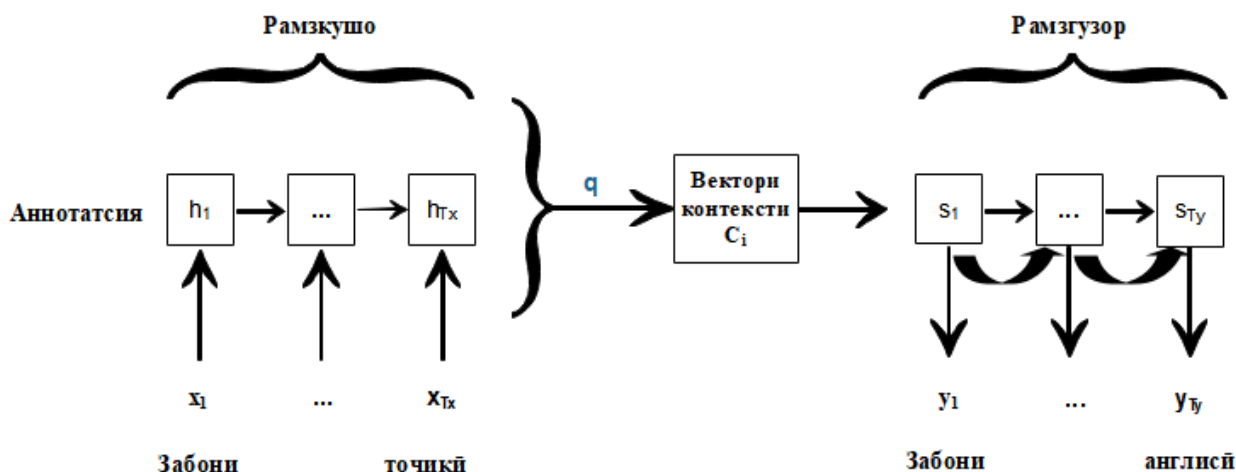
Агар низоми шинохтан дар хотираи худ садоҳои ҳарфҳои гуногунро дошта бошад, пас ин равиш, ки ба номбар кардани ҷузъҳои синф асос ёфтааст. Аммо танҳо дар ҳолатҳои, ки маълумоти воридшуда бо садои бегона, талаффуз рост меоянд, ҳарфҳои мувофиқ эътироф карда шаванд. Аз ин рӯ, ин усули нисбатан содда аст,

аммо он имкон медиҳад, ки низомҳои шинохти нутқ танҳо дар соҳаҳои муайяни татбиқе, ки пурра аз ӯҳдаи вазифаҳои онҳо баромада метавонанд, синтез карда шавад. Усули номбар кардани ҷузъҳои синф ба таври мақбул дар шароите амал мекунад, агар намунаи тасвирҳо ба тасвири истинод наздик бошад.



Расми 2 - Блок схемаи ворид кардани намуна ба манбаи додаҳо

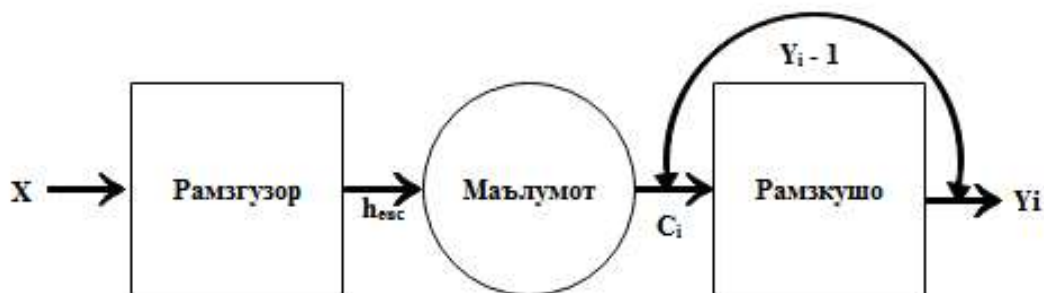
Дар § 1.4 алгоритмҳои шинохти нутқ дар асоси шабакаи нейронӣ ва татбиқи онҳо дар масъалаҳои истифодаи алгоритмҳо барои шинохти нутқ бо забонҳои англисӣ, русӣ баррасӣ шудаанд. Инчунин аз нав коркард намудани якчанд алгоритмҳо барои шинохти нутқ бо забони тоҷикӣ натиҷаҳо оварда шудаанд. Низомҳои шинохти нутқ шабакаи нейронӣ дар ҳадди аққал дар як қисми марҳилаи коркард шабакаҳои нейронҳои занҷириро истифода мебарад. Сохтори нисбатан бомуваффақияти шабакаҳои нейронҳои занҷирӣ барои шинохти нутқ ин занҷири дарози хотираи кӯтоҳмуддат мебошад ва яке аз усулҳои беҳтарин тавсифи моделҳо маҳсуб меёбад [5].



Расми 3 - Шабакаи нейронии рамзгузор-рамзкушо

Дар расми 3 намунаи тарчумаи мошин, яъне пайдарпаии дохилшавии маълумот ба забони тоҷикӣ ва содироти он бо забони англисӣ нишон дода шудааст. Муҳимият он аст, ки таваҷҷӯҳ ду таъсири мусбат дорад: якум, он кори шабакаро беҳтар мегардонад ва дуюм, рафтори шабакаҳои нейрониро, ки дарки он душвор аст, шарҳ медиҳад.

Дар расми 4 се модул дар модели TajNN-SAT нишон дода шуда, сохтори пирамидашакли он бо ду ҷузъи алоҳида тасвир шудааст. Чуноне, ки дар расм дида мешавад, рамзкушо хусусияти маҷбуркунии омӯзишро истифода мебарад. Ин маънои онро дорад, ки сигнали хориҷшавии дараҷаи қаблӣ ба қабати ҷорӣ рамзкушо ҳамчун воридшавӣ интиқол дода мешавад.



Расми 4 - Ҷузъҳои модели TajNN-SAT

Талафкунанда дар модели TajNN-SAT як пирамидашакли занҷири дарози хотираи кӯтоҳмуддати дучониба рамзгузори дохилшавӣ x ба функсияҳои сатҳи баланд h дида метавонад. Ҳамзамон рамзкушо шунаванда, ки ба таваҷҷӯҳ асос ёфтааст, аломатҳои y -ро аз h эҷод мекунад. Рамзгузор занҷири дарози хотираи кӯтоҳмуддати дучонибаро бо сохтори пирамидашакл истифода мебарад. Ин сохтори пирамидашакл аз рӯи дарозии h (намояндаи сатҳи баланди маълумоти дохилшавӣ x) кам мешавад. Бо сабаби раванди мураккаб доштани рамзкушо содироти додаҳои дахлдор аз миқдори зиёди қадамҳои муваққатии воридшавӣ аз ҷониби занҷири дарози хотираи кӯтоҳмуддати дучонибаи оддӣ хеле суст муттаҳид мешавад. Аз ин рӯ, истифодаи шакли пирамидаи занҷири дарози хотираи

кӯтоҳмуддати дучониба муфид аст, зеро он андозаи вақти рамзкушоиро кам мекунад, то ки ба камшавии ҳисоббарорӣ мураккаб оварад. Дар натиҷа, модели таваҷҷӯҳ метавонад маълумоти дахлдорро дар марҳилаҳои камтари вақт дарёфт кунад.

Дар **боби 2** “*Тарҳрезии амсилаҳо ва усулҳои шинохти нутқ*” натиҷаҳои асосии дар доираи диссертатсия ба даст оварда ба монанди амсилаи математикии симои рақамии нутқ, амсилаи компютери нишондиҳандаҳои нутқ бо забони тоҷикӣ ва усулҳои сохторбандии шинохти чараёни нутқ, усулҳои омӯзиши акустикӣ дар амсилаи TajNN-SAT дар мисоли нутқ бо забони тоҷикӣ баррасӣ карда шудаанд.

Дар § 2.1 *амсилаи математикии симои рақамии нутқ*, ки барои коркарди ибтидоии нутқ коркард карда шудааст, пешниҳод карда мешавад. Таҷриба нишон медиҳад, ки пас аз коркарди ибтидоии нутқ маҷмӯи фреймҳо ба калимаи муайян мувофиқ аст. Барои муқоиса кардани маълумоти ҳосилшуда ва маҷмӯи калимаҳои додашуда, муайян кардани хусусиятҳои овозӣ зарур аст. Дар асоси воҳидҳои муайяншуда масофаи байни векторҳои фреймҳои овозӣ ва векторҳои маҷмӯи қиматҳои инъикоси масъалаи шинохт ҳисоб карда мешавад ва вобаста ба наздикшавии векторҳо мувофиқат муайян карда мешавад.

Ба сифати метрика, масалан, он қимати миёнаи квадратири гирифтани мумкин аст, ки намуди зеринро дорад:

$$\rho = \sqrt{\frac{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2}{n}} \quad (1)$$

ки дар ин ҷо n – миқдори қиматҳои қабулшуда дар маҷмӯи фреймҳо, $x_i, y_i, i = \overline{1, n}$ – мутаносибан ҷузъҳои векторҳои соҳаи муайяни ва қиматҳои инъикоси масъалаи шинохт мебошанд. Чунин метрика аз сабаби калон будани қимати n ва на барои ҳама гуна қимати $i, i = \overline{1, n}$ мавҷуд будани маълумоти зарурии раванди шинохт натиҷаи дилхоҳро намедихад.

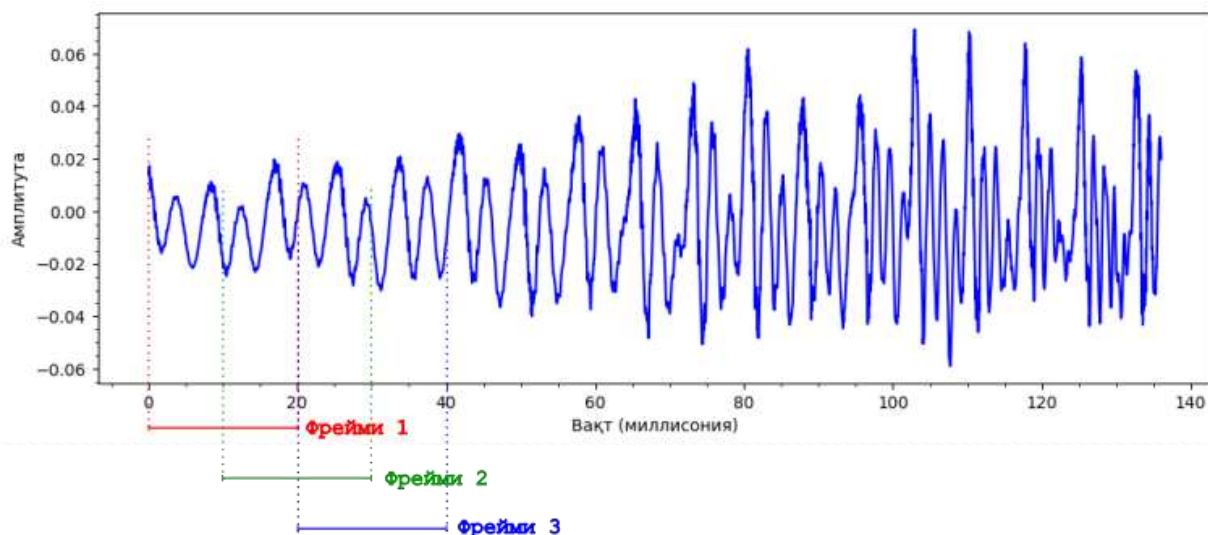
Нутқи инсон сигнали душвор ба ҳисоб меравад ва ин душворӣ аз рӯи мавҷуд будани миқдори зиёди хусусиятҳои нутқи инсон вобаста аст. Хусусиятҳои нутқи инсонро дар сатҳҳои гуногун, масалан, сатҳи акустикӣ, маъноӣ, забонӣ, равонӣ ва ғайра маънидод намудан мумкин аст. Ҳар як шахс дорои овози нодир аст ва ҳатто дар ҳолате, ки ҳамон як шахс ҳамон як калимаро талаффуз мекунад, натиҷаҳои ҳосилшаванда якхела нестанд. Бинобар ин, соҳаи таҳлили компютери нутқ самти афзалиятноки таҳқиқоти илмиро ташкил медиҳад.

Дар § 2.2 *усулҳои таҳлили хусусиятҳои воҳидҳои овозӣ* дар асоси натиҷаҳои таҳқиқотии хусусиятҳои воҳидҳои овозии бо забони тоҷикӣ талаффуз карда шуда, оварда шудаанд. Натиҷаҳои ҳосилшуда, вобаста ба муҳити барномаи компютери сабти овоз – ташкилкунандаи файли WAV, ки он ташкил карда шудааст, қиматҳои гуногуни амплитударо соҳибанд. Барои ягонакунии маълумоти ба даст овардашуда, ки қиматҳои ҳам мусбат ва ҳам манфиро дошта метавонанд, раванди муътадилкунии маълумот иҷро карда мешавад. Барои муътадилкунӣ маълумоти ҳосилшударо ба калонтарин қимати мутлақи амплитудаҳо тақсим кардан мумкин аст:

$$a_i = \frac{a_i}{\max(\text{abs}(\text{minVal}), \text{abs}(\text{maxVal}))}, \quad (2)$$

дар ин чо, a_i – қимати аз файл гирифташудаи амплитуда, функсияи $\max()$ – функсияе, ки қимати калонтарини аргументҳояшро бар мегардонад, функсияи $\text{abs}()$ – функсияе, ки қимати мутлақи аргументашро бар мегардонад, minVal ва maxVal – мувофиқан қиматҳои хурдтарин ва калонтарини амплитуда, ки аз файл гирифта шудаанд, мебошанд. Дар ин ҳолат, ҳамаи қиматҳои амплитуда аз порчаи $[-1; 1]$ қимат қабул мекунанд.

Қадами аввалини коркард пас аз гирифтани маълумот аз файл – ҷудокунии он ба фреймҳо аст. Фрейм – ин ҷудокунии маълумоти овозӣ ба фосилаҳои кӯтоҳмуддати вақт мебошад. Ҷудокунии маълумот ба фреймҳо бошад, пайдарпай иҷро нашуда, балки болои ҳам хобанда иҷро карда мешаванд. Яъне, охири як фрейм ба аввали фрейми дигар мувофиқ аст (расми 5).



Расми 5 - Ҷудокунии маълумоти овозӣ ба фреймҳо

Сабаби ҷудокунии нутқ ба фреймҳо аз он иборат аст, ки муқоисаро дар асоси фреймҳо нисбат ба муқоисакунӣ дар ҳар як лаҳзаи вақт иҷро кардан хеле осон аст. Болои ҳам хобидани аввал ва интиҳои фреймҳо имкон фароҳам меорад, ки раванди таҳлили фреймҳо суфта карда шавад.

Дар § 2.3 усулҳои сохторбандӣ ва шинохти ҷараёни нутқ ва истифодаи онҳо барои ҳалли масъалаи асосии рисолаи илмӣ, яъне шинохтагирии фармонҳо бо забони тоҷикӣ баррасӣ шудаанд.

Чун қоида, омӯзиши амиқи модели шабакаи нейронӣ боиси беҳтар шудани шинохти нутқ мегардад. Дар байни сохторҳои гуногун омӯзиши амиқ дар асоси алгоритми занҷири ибтидоии акустикаи омӯзишӣ барои шабакаи нейронӣ (ЗИАО-ШН) ва моделҳои зеҳни сунъӣ, ки дар амал натиҷаҳои назаррас доранд, бештар истифода мешаванд. Дар ин маврид, қолаби махсус ба худ модели акустикӣ, модели забонӣ ва механизми баробарнамоии овоз ва матнро якҷоя намуда, дар натиҷа низоми оддӣ шинохти худкори нутқро ташкил медиҳад. Ғайр аз он, омезиши қабатҳои шабакаи нейронҳои печон ва занҷири дарози хотираи кӯтоҳмуддат низ барои истихроҷи аломатҳо хуб кор мекунанд. Дар ин зерқисми диссертатсия

маълумоти муфассал оид ба қолабҳои ЗДИАО-ШН ва TajNN-SAT, ки дар меъмории замони муосир байни рамзгузор-рамзкушо маъмуланд, пешниҳод мегардад.

Дар § 2.4 усулҳои омӯзиши акустикӣ дар амсилаи TajNN-SAT хусусияти таҷрибаҳо ва натиҷаҳои онҳо дар муқоиса бо моделҳои истифодашавандаи нутқ мавриди муҳокима қарор дода мешаванд. Таҷрибаҳо барои дарёфти беҳтарин хусусиятҳо ва танзими онҳо барои модели TajNN-SAT анҷом дода мешаванд. Барои баргараф кардани маҳдудияти вақт ва зудтар ба даст овардани натиҷаҳо раванди таҳлил ва таҷрибаҳо бо маҷмӯи додаҳои Taj-Speech-Corpus иҷро карда шудаанд.

Дар асоси натиҷаи қорҳои таҳқиқотӣ дар диссертатсия маҷмӯи таҷрибаҳои ниҳой гузаронида шуданд. Барои баланд бардоштани самаранокӣ, ҳар ду модел TajNN-SAT ва ЗДИАО-ШН дар маҷмӯи маълумоти овозӣ бо забони тоҷикӣ, яъне дар Taj-Speech-Corpus мавриди омӯзиш қарор дода шуданд.

Ҷадвали 1 - Хусусиятҳои модели ЗДИАО-ШН барои маҷмӯи маълумоти Taj-Speech-Corpus

№	Муттаҳидшавии максималӣ	Қабати пурра пайвастшуда	Қабати занҷир	Дастрасии додаҳо	Басомади ҳатогиҳо дар рамзҳо %
1	Ҳа	Ҳа	1	Пайдарпай	0.81
2	Ҳа	Ҳа	1	Пайдарпай	0.81
3	Ҳа	Ҳа	5	Эҳтимолий	0.78
4	Ҳа	Ҳа	5	Эҳтимолий	0.80

Сарчашма: Таҳияи муаллиф.

Чуноне ки аз ҷадвали 1 дида мешавад, фоизи басомади ҳатогиҳо дар рамзҳо барои ҷалби таҳиягарон мебошад. Дар ин маврид, қабати занҷири дарозии хотираи кӯтоҳмуддати дучониба сатҳи максималии муттаҳидкунӣ ва пайвасти охири истифода мешавад, ки қабати охири ба ҳисоб меравад. Дар қабатҳои занҷири дарозии хотираи кӯтоҳмуддати дучониба андозаи пинҳонии ҳар як самт ба 256 баробар аст.

Ҷадвали 2 - Хусусиятҳои модели ЗДИАО-ШН дар маҷмӯи додаҳои соҳаҳои гуногун аз корпуси Taj-Speech-Corpus

№	Муттаҳидшавии максималӣ	Қабати пурра пайвастшуда	Қабати занҷир	Дастрасии додаҳо	Маҷмӯи додаҳо	Басомади ҳатогиҳо дар рамзҳо %
1	Ҳа	Ҳа	1	Пайдарпай	Тиб	0.85
2	Ҳа	Ҳа	1	Пайдарпай	Сайёҳӣ	0.80
3	Ҳа	Ҳа	1	Пайдарпай	Таълим	0.74
4	Ҳа	Ҳа	1	Пайдарпай	Афсонаҳои кӯдакона	0.78
5	Ҳа	Ҳа	5	Эҳтимолий	Тиб	0.98
6	Ҳа	Ҳа	5	Эҳтимолий	Сайёҳӣ	0.80
7	Ҳа	Ҳа	5	Эҳтимолий	Таълим	0.76
8	Ҳа	Ҳа	5	Эҳтимолий	Афсонаҳои кӯдакона	0.86

Сарчашма: Таҳияи муаллиф.

Чуноне ки дар ҷадвали 2 дида мешавад, натиҷаи сохтории ЗДИАО-ШН барои маҷмӯи додаҳои Taj-Speech-Corpus нисбат ба натиҷа барои маҷмӯи маълумоти

соҳати тиб начандон хуб аст. Инро бо таъсири шумораи сифрҳо, ки ҳангоми истифодаи калимаҳои соҳавии тиб дар Taj-Speech-Corpus ба шабака интиқол дода мешаванд, шарҳ додан мумкин аст. Азбаски дарозии намунаҳо бо пуррагӣ дар Taj-Speech-Corpus гуногун мебошанд, барои яхела кардани андозаи додаҳои воридшаванда сифрҳо ба намунаҳои кӯтоҳтар илова карда мешаванд, то ки ҳамаи додаҳои воридшаванда мутобиқи дарозӣ баробар шаванд.

Дар умум, на танҳо моделҳои асосӣ, ки дар он усулҳои прогрессивии таҳлил ва синтези нутқ таҳлил карда мешаванд, балки баъзе моделҳои, ки дар айнаи ҳол мустақиланд, истифода мешаванд.

Дар **боби 3** “Маҷмӯи барномаҳои компютери TAJIK-AVCR” амсиласозии компютери таҳияи низомҳои коркарди худкори иттилоот бо забони табиӣ бо назардошти равиши ба объект нигаронидашуда баррасӣ карда шудааст. Дар асоси маҷмӯи диаграммаҳои забони UML тарҳи намунавии низоми иттилоотии коркарди маълумоти матнӣ бо забони тоҷикӣ бо назардошти амсилаи амал, амали мутақобила, сохтор ва воқеият таҳия карда шуд. Инчунин натиҷаҳои истифодаи амалии низоми шинохтагирии нутқ ва баҳодиҳии самаранокии он дар доираи васеи корбарон пешниҳод карда мешаванд.

Дар § 3.1 амсилаи математикии равандҳои корӣ барои лоиҳаи шинохти нутқ бо забони тоҷикӣ пешниҳод карда шудаанд. Маҷмӯи барномаҳои компютери TAJIK-AVCR лоиҳаҳои ҳисоббарор мебошанд, ки барои нутқи муттасили корбарро аз чараёни умумии садоҳо фарқ кардан имконият медиҳанд. Ин низомҳо ба технологияи шинохти нутқ асос ёфтаанд, ки тавассути раванди шинохти калимаҳо аз овоз ба сигналҳои рақамӣ табдил меёбанд. Ҳамчунин роҳи ҳалли масъалаи ин коркарди модулҳои таъминоти барномавӣ барои шинохти нутқ ва ҷустуҷӯи додаҳои зарурӣ дар матн ҳосилшуда, пешниҳод карда мешаванд [6].

Барои оғоз, ташкили сохтори низом бояд таъсис дода шавад. Дар доираи ин низом унсурҳои зерин муҳим хоҳанд буд:

1. *Равандҳои корӣ* – *P*: бақайдгирӣ ва назорати корбарон; ташкили манбаи додаҳои луғат; кор бо сабти овоз; шинохти нутқ; таҳлили натиҷаи шинохти нутқ.

2. *Субъектҳо* – *S*: маъмур; корбари оддӣ; забоншинос (мутахассис); муҳандис.

3. *Функсияҳои имконпазир* – *F*: бақайдгирии корбар; аз қайд гузаштани корбар; воридшавӣ ба низом; таҳрири “мизи кории” корбар; ворид кардани элементи матнӣ ба луғат; таҳрир кардани элементи матнӣ дар луғат; сабти овоз ва коркарди аввали он; тақсим намудани сабт ба воҳидҳои (қолаб) овозӣ; идоракунии воҳидҳои (қолаб) овозӣ; ташҳиси овози воридшуда; муқоиса бо воҳидҳои матнии дар луғат ҷойдошта; ҳисоби фоизи дақиқи шинохти нутқ; пешниҳоди воҳиди матнӣ мувофиқи натиҷаи шинохти нутқ; таъмини амнияти иттилоот.

4. *Пояи додаҳо ва ҷадвалҳои он* – *T*: корбари низоми иттилоотӣ: рақами тартибии корбар, ном, насаб, чинс, санаи таваллуд, почтаи электронӣ, маълумот оид ба таҳсилот, вазифа; элементи матн: рақами тартибии элементи матн, матн, намуд (ҳарф, ҳиҷо, калима, ибора, ҷумла), санаи ворид кардан, соҳаи предметӣ, эзоҳот; элементи овоз: рақами тартибии элементи овоз, сабти овоз, намуд (wav, mp3), санаи сабт кардан, рақами тартибии корбар, эзоҳот; шинохти нутқ: рақами тартибии элементи матн, рақами тартибии элементи овоз, фоизи дақиқи шинохт, ҳолати шинохт, эзоҳот; натиҷаи таҳлили фаъолнокии корбар: рақами тартибии корбар, миқдори воҳиди овозии сабткарда, фоизи миёнаи шинохти нутқ.

Моделҳои математикӣ муносибати байни субъектҳо, функсияҳо ва маълумотро дар таҳлили омӯри тавсиф мекунад:

$$I_{S_i}^{P_j} \Rightarrow f_k(t_1, t_2, t_3, \dots, t_q) \Rightarrow I_n \quad (3)$$

дар ин ҷо:

$I_{S_i}^{P_j}$ – иштироки субъекти $S(i)$ дар раванди $P(j)$ дар муҳити кории I , $i=1,2,\dots,4$ ва $j=1,2,\dots,5$;

$f_k(t_1, t_2, t_3, \dots, t_q)$ – дастрасии функсияи $F(k)$ барои коркарди додаҳои ҷадвали $T(q)$, $k=1,2,\dots,14$ $q=1,2,\dots,5$.

I_n – пешниҳоди натиҷа дар муҳити кории нав.

Моделҳои математикиро барои иҷрои функсияи f_1 чунин амалӣ кардан мумкин аст. Субъекти S_1 (*маъмури* низом) аз раванди P_1 (*бақайдгирӣ ва назорати корбарон*) ва функсияи f_1 (*бақайдгирии корбар*) бо истифодаи додаҳои ҷадвалҳои t_1 (*корбарони низоми иттилоотӣ*) ба роҳ менамояд. Дар натиҷаи ин кор, маълумоти нави $I(n)$ ҳосил мешавад, ки он барои шинохти овоз равона карда мешавад.

Намунаи амсилаи математикӣ бо истифода аз ин маълумот чунин аст:

$$I_{(S_1)}^{(P_1)} \Rightarrow f_1(t_1) \Rightarrow I_n \quad (4)$$

дар ин ҷо:

$I_{(S_1)}^{(P_1)}$ – муҳити корӣ барои бақайдгирии корбар;

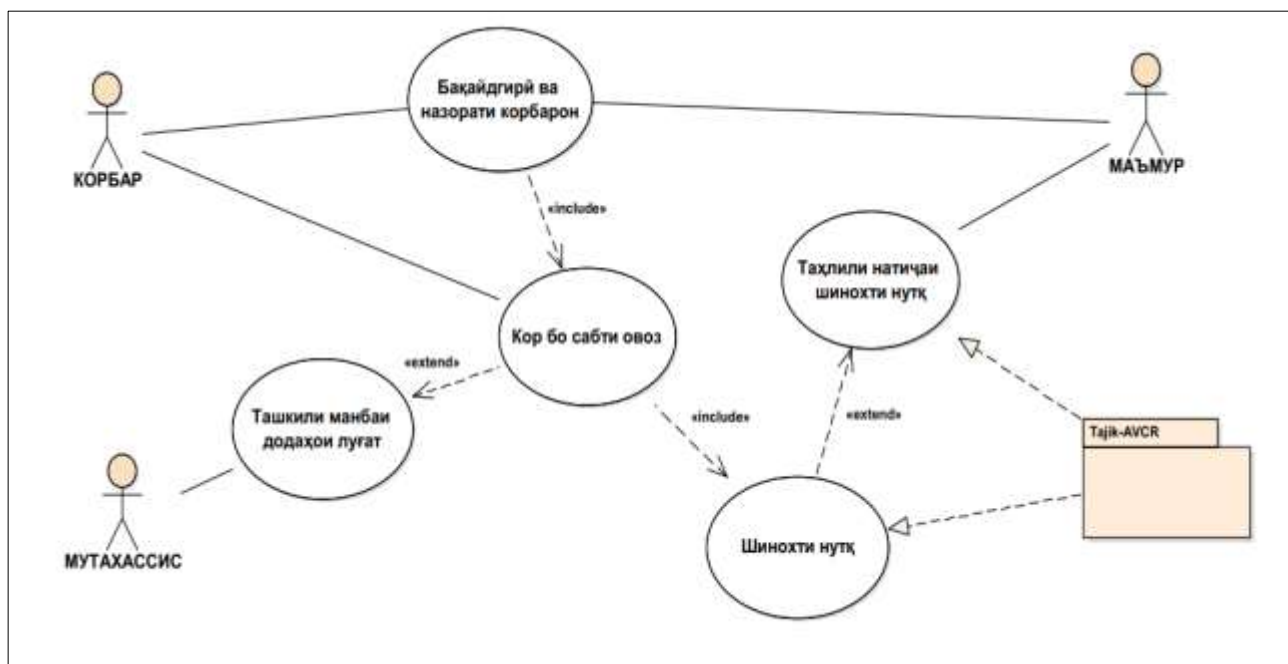
$f_1(t_1)$ – бақайдгирии корбар ва нигоҳдории маълумот дар ҷадвали корбари низоми иттилоотӣ;

I_n – маълумоти нав оид ба сабти овоз ва ҳолати он.

Дар § 3.2 *сохтори мантиқии низоми шинохтагирии овоз Tajik-AVCR* масъалаи амсиласозии компютерӣ ва пешниҳоди лоиҳаи таъминоти барномавӣ тавассути истифодаи васеи CASE-воситаҳо ҳалли худро меёбанд. Дар ин самт талаботи система, сохтори мантиқӣ, ташхиси объектҳо, муносибати мантиқии байни онҳо, равандҳои корӣ ва амсилаи воқеӣ бо истифода аз воситаҳои махсус муайян ва тасвир карда мешаванд. Марҳилаи амсиласозӣ аз муайян кардани маҷмӯи диаграммаҳо бо истифодаи забони UML барои таҳлили талабот ва пешниҳод ба марҳилаи коркард дар вақти воқеӣ муҳиманд, иборат аст [7].

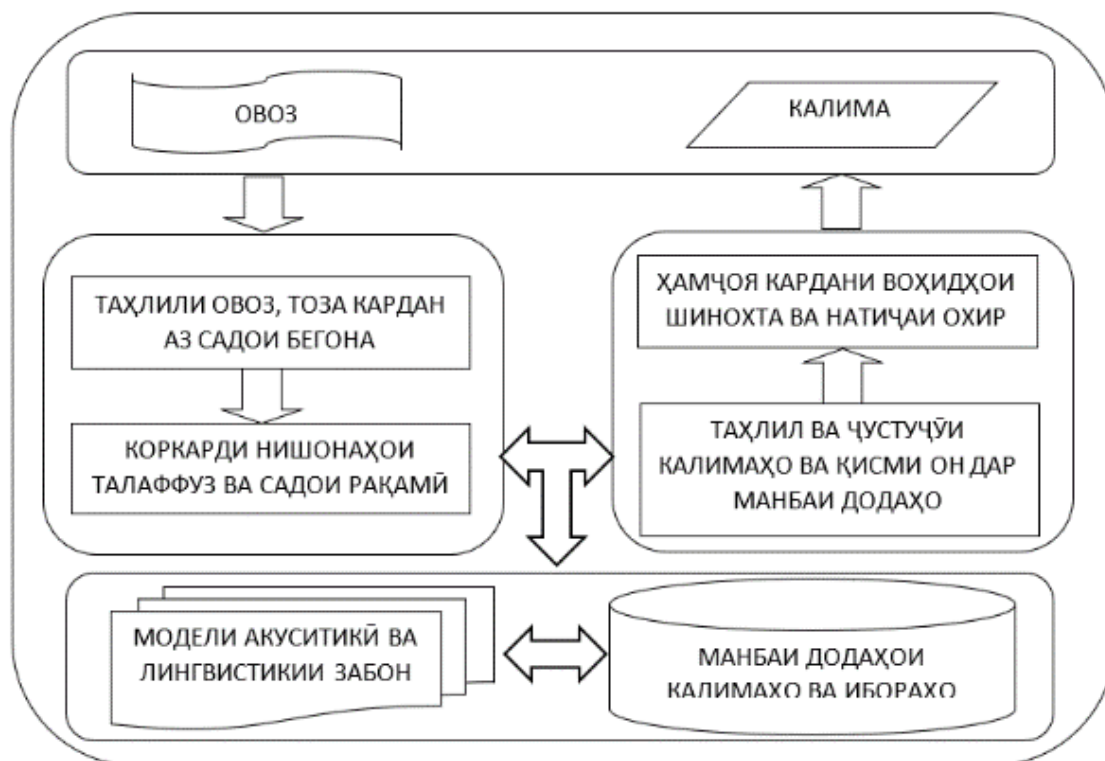
Диаграммаҳои асосие, ки бояд барои лоиҳаи Tajik-AVCR муайян карда дар рӯйхати зерин оварда мешавад. Дар умум, низоми таҳияшудаи Tajik-AVCR ду омил дорад. Дар аввал, шахсони иштироккунанда нишон дода мешаванд, ки ҳамчун корбари он ба ҳисоб мераванд. Омили дуюм ин муайян кардани ҳамаи ҳолатҳои истифодабари дар раванди шинохтагирии фармоиши овоз, ки аз як тараф мониторинги амсилаи рафтори низоми иттилоотиро таъриф менамояд. Ин қадами асосӣ дар раванди амсиласозии компютери низомҳои иттилоотӣ ба ҳисоб меравад.

Амсилаи рафтори низоми Tajik-AVCR дар асоси диаграммаи вазъияти корфармои дар расми 6 оварда шудааст.



Расми 6 - Амсилаи рафтори низоми Tajik-AVCR

Дар § 3.3 истифодаи амалии низоми шинохтагирии овозӣ дар комплекси Tajik-AVCR бо пешниҳоди сохтори мантиқии низоми шинохти овоз ва интерфейси корбар мавриди муҳокима қарор дода шудааст.

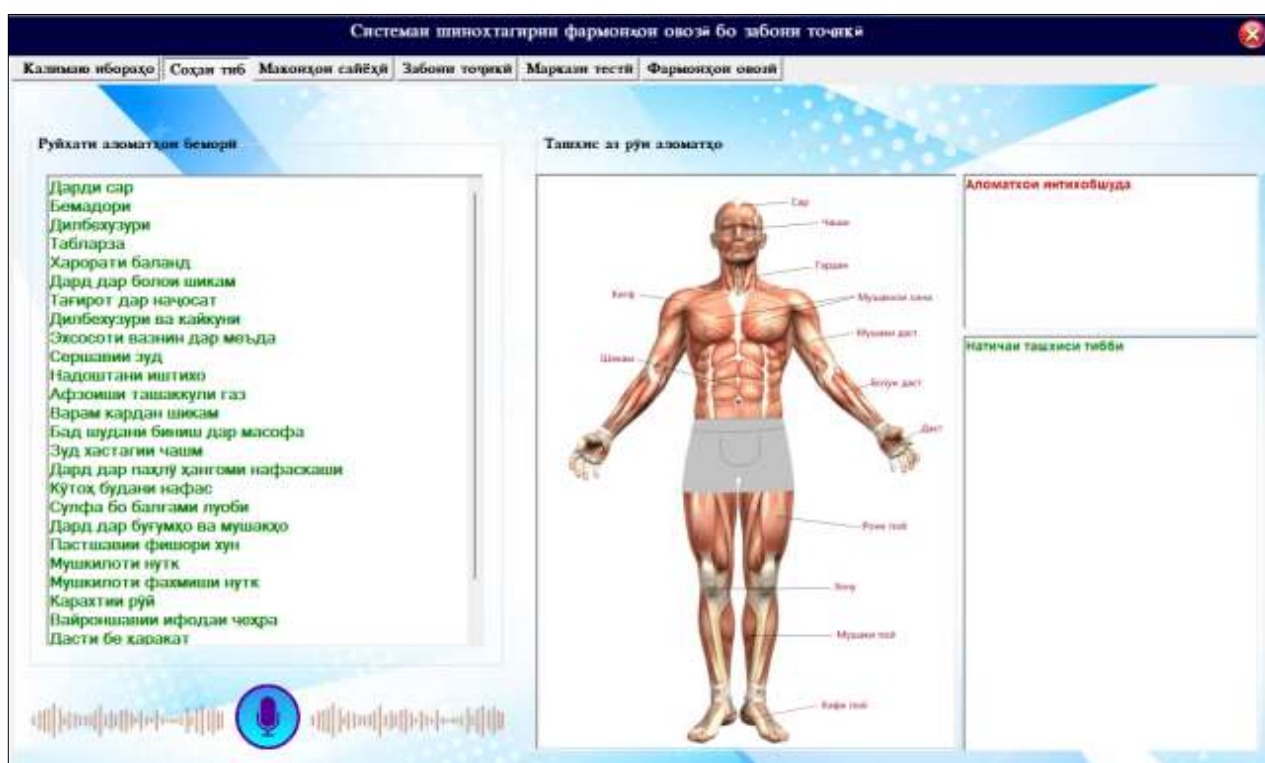


Расми 7 - Сохтори мантиқии низоми шинохти овоз – Tajik-AVCR

Сохтори мантиқии низоми шинохти овоз. Диаграммаи кори низоми Tajik-AVCR дар расми 7 нишон дода шудааст. Дар даромади барнома фармони овозӣ

қабул карда мешавад, ки шумо мехоҳед ҷустуҷӯ кунед. Дар аввал, коркарди алоҳидаи ин маълумот ба анҷом дода мешавад. Сипас, масъалаи коркарди хосиятҳои нутқи талаффузшуда дар намуди садои рақамӣ ба роҳ монда мешавад. Дар мавриди коркарди нутқ амсилаи акустикӣ ва лингвистикӣ, ки дар асоси амсилаи забони тоҷикӣ омода шудаанд, ғаёб карда мешаванд.

Пас аз коркарди нутқи воридшуда таъминоти барномавӣ матни шинохта ва ҷустуҷӯ шударо дар қисми махсуси оинаи матнӣ намоиш медиҳад. Барои ғаёб кардани имконияти пурраи низоми шинохт таъминоти барномаи ба танзим оварда мешавад. Мавҷудияти модулҳои иловагии барномавӣ дар компютери корбар яке аз шартҳои самараноки он ба ҳисоб меравад. Барнома бо истифода аз файли корӣ оғоз карда мешавад Tajik-AVCR.exe. Пас аз оғоз, интерфейси корбар дастрас мешавад (расми 8).



Расми 8 - Интерфейси корбари Tajik-AVCR дар соҳаи тиб

Интерфейси корбар барои ҳамкориҳои интерактивии корбар бо таъминоти барномавӣ дар асоси барномарезии ба объект нигаронидашуда, таҳия шудааст. Дар ин мақола, интерфейси корбар дар шакли соддакардашуда пешниҳод карда шудааст, ки ҳангоми амалисозии таъминоти барномавӣ якҷанд тағйиротҳоро талаб мекунад.

Дар § 3.4 баҳодиҳии самаранокии низоми шинохтагирии нутқ дар асоси натиҷаҳои дар истифодаи амалӣ ба даст оварда шуда, вобаст аз якҷанд хусусиятҳои корбарон сифат ва самарои шинохти нутқ дар маҷмӯи барномаҳои Tajik-AVCR пешниҳод карда шудаанд. Ҳаминро қайд кардан зарур аст, ки байни сӣ нафар корбарон фоизи миёнаи шинохти нутқ 76,48% таъхис медиҳад. Фосилаи фоизи шинохти нутқ аз 64,33 то 85,47 фоизро таъхис медиҳад. Натиҷаи ба даст оварда, албатта дар шароити имрӯзи коркарди низомҳои худкори шинохти нутқ ва бозори таъминоти барномавӣ ба талабот ҷавобгӯ мебошад. Илова бар ин, корбарон аз рӯйи

чинс ба ду гуруҳ мардҳо ва занҳо ҷудо карда шуданд. Инчунин барои таҳлили мушаххас корбарон аз рӯи сини сол ба се гуруҳ тақсим карда шуданд: аз 14 то 24 сола, аз 25 то 44 сола ва аз 45 то 60 сола.

Натиҷаҳои мавҷударо таҳлил намуда, ба хулосае омадан мумкин аст, ки ҳангоми таҳлили шинохти нутқ ва пешниҳоди матн, фарқи миёнаи дурусти шинохт байни мард ва зан 2,36 фоизро ташкил медиҳад. Дар ҳоле, ки таҳлил бо истифода аз гуруҳбандӣ ба сини сол нишон медиҳад, ки матнҳои аз якдигар гуногун талаффуз шуда, ба дақиқии шинохти миёнаи 76,74% баробар аст. Аз тарафи корбарон рӯйхати муайяни калима ва ибораҳои дар луғати низоми шинохт нигоҳ дошта шуда, талаффуз карда шуданд. Бояд қайд кард, ки пайдарпаии талаффуз ба инобат гирифта намешавад ва танҳои фоизи шинохти нутқ ба назар гирифта мешавад. Дар расми 9 фоизи шинохти нутқ вобаста аз чинс ва синну соли корбарон, ба намуди гистограмма оварда шудааст.



Расми 9 - Фоизи шинохти нутқ дар низоми Tajik-AVCR вобаста аз чинс ва синну сол

Дар ин маврид, қимати миёнаи дақиқии шинохти нутқ байни занҳои синну солашон аз 25 то 44 солаи зиёдтар ба назар мерасад. Аммо дақиқии шинохти нутқ байни мардҳои синну солашон аз 14 то 24 сола нисбатан фоизи камро соҳиб гашт. Ин муаммо ба раванди шинохти нутқ вобаста буда, усули талаффузи табиӣ корбаронро ифода менамояд.

Аз рӯи натиҷаҳои таҳқиқот, метавон ба хулосае омад, ки барои ошкор кардани фармонҳои овозӣ дар талаффузи давомдор усули истифодаи алгоритмҳои шабаки нейронӣ эҳтимолияти қобили қабул ва самарабахш ба ҳисоб меравад. Инчунин маълум гашт, ки эҳтимолияти дуруст ҷустуҷӯ кардани фармони овозӣ дар матни талаффуз шуда, танҳо ба сифати шинохтан ва дар матн мавҷудияти кофии маълумоти ба соҳаи муайян вобаста аст. Дар умум натиҷа нишон медиҳад, ки низоми шинохти коркардшуда, самаранок буда, вазифаҳои гузошташударо мувофиқи талаботи имрӯза иҷро мекунад.

ХУЛОСА

1. Дар заминаи таҳлили таҳқиқоти илмӣ ва амалӣ дар соҳаи забоншиносии компютерӣ ва технологияҳои овозӣ дар Ҷумҳурии Тоҷикистон ва берун аз он, таҳқиқоти назариявӣ барои таҳқиқоти илмӣ **муқаррар гардиданд**, ки аз таҳияи амсилаҳои математикӣ ва компютерӣ, усулҳои коркарди нутқ дар асоси шабакаҳои нейронӣ **омӯхта шудаанд** ва татбиқи онҳо дар низоми худкори шинохти нутқи тоҷикӣ иборат мебошад [1-М], [3-М], [7-М], [9-М].

2. Бо мақсади ҳалли масъалаи коркарди низоми иттилоотии шинохтагирии фармонҳои овозӣ бо забони тоҷикӣ усулҳои нав дар асоси шабакаҳои нейронӣ **таҳия, аз нав коркард ва пешниҳод карда шудаанд** [2-М], [10-М], [12-М].

3. Амсилаҳои математикӣ ва компютерӣ барои симои рақамии нутқ **таҳия гардида**, усулҳои коркарди воҳидҳои овозӣ **коркард** ва дар намуди алгоритмҳои алоҳида **татбиқ карда шуданд** [2-М], [4-М], [17-М].

4. Усулҳои маъмули бо шабакаи нейронӣ вобаста ба монанди CNN-BLSTM ва TajNN-SAT **пешниҳод шудаанд**, ки аз комплекси амсилаҳо ва алгоритмҳо дар масъалаҳои шинохтагирии воҳидҳои овозӣ бо забони табиӣ **татбиқ карда шуданд** [2-М], [4-М], [8-М], [11-М], [14-М], [15-М].

5. Дар рисолаи илмӣ масъалаҳои асосии ҷамъоварӣ, нигоҳдорӣ, коркард, ва рақамигардонии воҳидҳои овозӣ бо забони тоҷикӣ **муқаррар шуда**, симои рақамии онҳо **пешниҳод карда шудаанд** [5-М], [7-М], [12-М], [13-М], [17-М], [18-М].

6. Маҷмӯи барномаҳои компютери TAJK-AVCR **коркард карда шудааст**, ки самаранокии кори диссертатсионӣ, яъне амсилаҳои математикӣ ва компютерӣ, усулҳои махсусгардонидаи шабакаҳои нейронӣ, алгоритмҳои коркарди воҳидҳои овозӣ **асоснок карда шудаанд** [4-М], [6-М], [16-М].

7. Натиҷаҳои илмӣ-таҳқиқотии кори диссертатсионӣ барои омӯзиши хусусиятҳои забони тоҷикӣ барои ҳамаи қорбарони технологияҳои иттилоотӣ заминаи устувор **ба ҳисоб меравад** ва онҳо дар намуди комплекси барномаҳои TAJK-AVCR дар соҳаи илм ва маориф дар Донишгоҳи давлатии ҳуқуқ, бизнес ва сиёсати Тоҷикистон ва Донишкадаи саноат ва хизматрасонӣ, аз рӯи соҳаи тиб дар ҶДММ “Баҳри Шифо” ва ҶДММ “Ал-Финур”, аз рӯи соҳаи сайёҳӣ дар ҶДММ “Марвел Вита Травел” ва ширкати сайёҳии “Панҷистон” дар **истифода қарор дода шудаанд**, ки тасдиқи амалии натиҷаҳоро **асоснок мекунанд** [2-М], [5-М], [15-М].

ТАВСИЯҲО

Натиҷаҳои асосии илмӣ диссертатсия ва татбиқи амалии онҳо рушди минбаъдаи назарияи тарҳрезии пойгоҳи додаҳо ва ҳалли масъалаи илмӣ, яъне таҳияи усулҳои самараноки ташаккулдиҳии низомҳои иттилоотӣ ва замимаҳои шинохти нутқ, ки дар иҷрои вазифаҳои иттилоотӣ ва техникӣ дар муассисаҳои таҳсилоти олии касбӣ, махсусан дар раванди таълим, тандурустӣ, сайёҳӣ ва ғайра аҳамияти амалӣ доранд, таъмин менамоянд.

Аҳамияти амалии таҳқиқоти диссертатсионӣ аз таҳияи методикаи умумии моделиронии компютери пойгоҳи додаҳои шинохтагирии нутқ иборат аст. Воридоти тағйироти муайян ба алгоритм ва системаи барномавӣ имконият медиҳад, ки дар қолаби амсилаи мавҷуда тарҳи амсилаҳои муассисаҳои таҳсилоти дигар рехта шавад. Технологияи ҳосилшуда метавонад ба болоравии амалияи амсиласозии компютерӣ саҳми арзанда гузорад.

РҶҶҲАТИ АДАБИЁТИ ИСТИФОДАШУДА

[1]. Паёми Президенти Ҷумҳурии Тоҷикистон мухтарам Эмомалӣ Раҳмон дар бораи самтҳои асосии сиёсати дохилӣ ва хориҷии ҷумҳурӣ. 28.12.2024. www.president.tj.

[2]. Стратегияи миллии рушди Ҷумҳурии Тоҷикистон барои давраи то соли 2030. Қарори Маҷлиси намояндагони Маҷлиси Олии Ҷумҳурии Тоҷикистон аз 1 декабри соли 2016, № 636

[3]. Барский, А.Б. Искусственный интеллект и интеллектуальные системы управления: монография / А.Б. Барский. - Москва: РУСАЙНС, 2022.- 185 с.

[4]. Косимов, А.А. Становление компьютерной лингвистики Таджикистана / А. А. Косимов. – Душанбе: Ирфон, 2021. – 102 с. – ISBN 978-99975-63-74-3. – EDN GBKWZO.

[5]. Белоруцкий, Р. Ю. Распознавание речи на основе сверточных нейронных сетей / Р.Ю. Белоруцкий, С.В. Житник. //Вопросы радиоэлектроники. – 2019. – №. 4. – С. 47-52.

[6]. Усманов, З. Д. Низомҳои худкори коркарди маълумот бо забони тоҷикӣ / З. Д. Усманов, Х. А. Худойбердиев. – Худжанд: Ирфон, 2022. – 186 с. – ISBN 978-99985—6624-8. – EDN ULPYYQ.

[7]. Худойбердиев, Х.А. Моделирование системы автоматической обработки текста на таджикском языке / Х. А. Худойбердиев // International Journal of Open Information Technologies. – 2023. – Т. 11, № 3. – С. 27-33. – EDN KRBOBH.

РҶҶҲАТИ ИНТИШОРОТИ ИЛМИИ ДОВТАЛАБИ ДАРАҶАИ ИЛМӢ

*Мақолаҳо дар маҷалла ва нашрияҳои илмӣ, ки аз тарафи КОА
назди Президенти Ҷумҳурии Тоҷикистон*

[1-М] Дадобоев, А.А. Модели математикии коркарди овоз ва ҳосиятҳои иттилоотикунонӣ [Матн] / А.А. Дадобоев // Номаи донишгоҳ. Силсилаи илмҳои табиатшиносӣ ва иқтисодӣ. -Хучанд. - 2023 -№1 (64) - С. 12-15. ISSN 2077-4974.

[2-М] Дадобоев, А.А. Модели математикии шинохтагирии овоз [Матн] / А.А. Дадобоев / Паёми политехникӣ: Бахши Интеллект. Инноватсия. Инвеститсия – 2023 -№ (61) –С. 27-33. ISSN 2520-2235.

[3-М] Дадобоев, А.А. Алгоритми бозшиносии нутқ [Матн] /А.А. Дадобоев // Номаи донишгоҳ. Силсилаи илмҳои табиатшиносӣ ва иқтисодӣ. Хучанд. – 2024- №1 (68) –С. 35 – 40. ISSN 2077-4974.

[4-М] Дадобоев, А.А. Таъминоти барномавӣ ва ҷудокунии нишонаҳои овозии нутқ [Матн] / А.А. Дадобоев / Паёми политехникӣ: Бахши Интеллект. Инноватсия. Инвеститсия – 2024 -№ (64) –С. 53-57. ISSN 2520-2235.

[5-М] Дадобоев, А.А. Низомҳои ҷустуҷӯ ва шинохти фармони овозӣ дар нутқи сарбастии забони тоҷикӣ [Матн] / Х.А. Худойбердиев, А.А. Дадобоев, Ш.Н. Ашурова // Паёми ДТТ – 2024. – № 3 (58). – С. 181 – 189.

[6-М] Анварзода А.А. Хусусиятҳои асосии системаҳои мобилӣ ва барномаҳои кроссплатформаи барои шинохти нутқ [Матн] / Б.Х. Ашурзода, А.С. Бобоев, А.А. Анварзода // Паёми ДТТ – 2025. – № 3 (71). – С. 63 – 67.

Мақолаҳо дар дигар маҷаллаҳои илмӣ

[7-М] Дадобоев, А.А. Усулҳои шинохтагирии нутқ ва таснифоти он [Матн] / Ш.М. Исмоилов, А.А. Дадобоев // Маҷлаи конференсияи байналмилалӣ илмӣ “Муаммоҳо ва

перспективаи алоқамандӣ дар ҷаҳони муосир», «Илми муосир», ш. Петрозаводск, 2021. – С. 166-173.

[8-М] Дадобоев, А.А. Барномаҳои компютерӣ барои таъминоти шинохтагии нутқ ва сегментатсияи нутқ [Матн] / А.А. Дадобоев // Маводи конференсияи байналмилалӣ илмӣ “Илмҳои”, «Илми муосир», ш. Петрозаводск, 2022. – С. 21 - 25.

[9-М] Дадобоев, А.А. Таснифоти назариявии модели шинохтагии овоз [Матн] / А.А. Дадобоев // Маводи конференсияи байналмилалӣ илмӣ “Пешсафони ИДМ”, ш. Тошканд, 2023. – С. 249 - 253.

[10-М] Дадобоев, А.А. Усулҳои асосии ҳалли масъалаҳои шинохтагии нутқ [Матн] / А.А. Дадобоев // Маводи конференсияи байналмилалӣ илмӣ “Муҳаққиқи ҷавони ИДМ”, ш. Тошканд, 2024. – С. 638 - 643.

[11-М] Дадобоев, А.А. Методы и алгоритмы распознавания речи [Матн] / А.А. Дадобоев // Маводи конференсияи илмӣ-амалии профессорон, омӯзгорон ва муҳаққиқони ҷавони ДДХБСТ “Илм ва инноватсия баҳри иҷроии ҳадафҳои миллӣ”, ш. Хучанд, “Дабир”, 2019.-С. 111-114.

[12-М] Дадобоев, А.А. Таснифи системаҳои шинохтагии овоз [Матн] / А.А. Дадобоев // Маводи конфронси илмӣ-амалии профессорону омӯзгорон ва муҳаққиқони ҷавон “Ҳамгироии илм ва истеҳсолот баҳри татбиқи ҳадафҳои стратегии миллӣ”.-Хучанд: “Дабир”, 2022.- С. 144-147.

[13-М] Дадобоев, А.А. Таснифоти назариявии модели шинохтагии овоз [Матн] / Ҳ.Ш. Ғаюров, А.А. Дадобоев // Маводи конфронси илмӣ-амалии ҷумҳуриявӣ “Рисолати илми физика дар инкишофи техника ва технологияи муосир”.-Хучанд: “Дабир”, 2023.- С. 113-117.

[14-М] Дадобоев, А.А. Таснифи системаҳои шинохтагии овоз [Матн] / А.А. Дадобоев // Маводи конфронси илмӣ-амалии профессорону омӯзгорон ва муҳаққиқони ҷавон “Рушди илм ва инноватсия дар шароити рақамикунонӣ”.-Хучанд: “Дабир”, 2023.- С. 153-158.

[15-М] Анварзода А.А. Методҳо ва алгоритмҳои синтези нутқ дар асоси матн. [Матн] / Худойбердиев Х.А., Ашурзода Б.Х., Анварзода А.А., Ашурова Ш.Н. // Паёми ДПДТТ ба номи М.С.Осими. – 2025. – № 3(35).

Шаҳодатномаҳои муаллифӣ ва бақайдгирӣ давлатии захираҳои иттилоотӣ

[16-М] Дадобоев, А.А. Web-приложение “Системаи шинохтагии фармонҳои тоҷикӣ” [SOFT] / А.А. Дадобоев // – 23.02.2024. – № 188.

[17-М] Дадобоев, А.А. Web-приложение “Системаи шинохтагии овоз бо забони тоҷикӣ” [SOFT] / А.А. Дадобоев, Х.А. Худойбердиев, Д.Ф.Маҳкамов // – 10.07.2024. – № 4202400519.

[18-М] Дадобоев, А.А. Web-приложение “Web-приложение taj-speech-corpus – Формирование речевых баз данных для таджикского языкового корпуса” [SOFT] / А.Т.Мақсуди, Х.А. Худойбердиев, А.А. Дадобоев, Ш.Н. Ашурова, А.А. Назаров, Д.Ф. Маҳкамов // – 17.09.2024. – № 4202400521.

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РЕСПУБЛИКИ ТАДЖИКИСТАН
ТАДЖИКСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ПРАВА,
БИЗНЕСА И ПОЛИТИКИ**

УДК: 811.21; 004.9

На правах рукописи



АНВАРЗОДА АКМАЛ АНВАР

**ПРОБЛЕМЫ РАСПОЗНАВАНИЯ ГОЛОСОВЫХ КОМАНД И
ИХ ПРИМЕНЕНИЕ В КОМПЬЮТЕРНОЙ МОДЕЛИ**

АВТОРЕФЕРАТ

диссертации на соискание ученой степени доктора философии (PhD) - доктор по специальности 6D070400 – Вычислительная техника и программное обеспечение (6D070402 – Вычислительные машины, комплексы и компьютерные сети)

Худжанд – 2025

Диссертация выполнена на кафедре информационно-коммуникационных технологий и программирования Таджикского государственного университета права, бизнеса и политики

Научный руководитель: **Худойбердиев Хуршед Атохонович** - доктор технических наук, доцент кафедры программирования и информационных систем политехнического института Таджикского технического университета имени академика М.С.Осими в городе Худжанд

Официальные оппоненты: **Рахимов Нодир Одилович** – доктор технических наук, доцент кафедры программное обеспечение информационных технологий Ташкентского государственного университета информационных технологий имени Мухаммада аль-Хорезми

Саидзода Исроил Махмад - кандидат технических наук, доцент, декан механико-математического факультета Национального университета Таджикистана

Ведущее организация: Технологический университет Таджикистана

Защита диссертации состоится «27» февраля в 2025 году в 14:00 на заседании Диссертационного совета 6Д.КОА-049 при Техническом университете Таджикистана имени академика М.С.Осими, по адресу 734042, г. Душанбе, проспект Академиков Раджабовых, 10А.

С диссертацией и авторефератом можно ознакомиться в библиотеке и на официальном сайте Таджикского технического университета имени академика М.С. Осими <http://web.ttu.tj>.

Автореферат разослан «__» _____ 2025 г.

Ученый секретарь
диссертационного совета,
кандидат технических наук, доцент



Султонзода Ш.М.

ВВЕДЕНИЕ

Актуальность темы исследования. Для достижения стратегических целей страны в Республике Таджикистан включены ряд актуальных вопросов о внедрении современных информационных технологий. Данный вопрос является одним из важных аспектов развития образования, а также проведения исследовательских работ в области естественных, технических и точных наук.

В своем послании основоположник национального мира и единства – Лидер нации, Президент Республики Таджикистан, уважаемый Эмомали Рахмон отметил важность развития информатики, цифровизации, внедрения современных технологий во всех сферах нашего общества (28 декабря 2024 года) [1].

Развитие мышления подростков и молодежи непосредственно привело к возникновению информационной культуры современной эпохи, постепенно приближающейся к единому информационному пространству. В современном мире развитие науки и техники тесно связано с эволюцией средств информационных и цифровых технологий. Сегодня большую популярность приобретают интерактивные и диалоговые интерфейсы, в которых реализованы мультимедийное взаимодействие человека и машины. Проведения исследований в сфере речевых технологий является наиболее из актуальных проблем для реализации общения человека с компьютерными средствами. Распознавание речи и голосовых команд можно считать, как важной частью автоматизированных систем с человеко-машинным речевым интерфейсом, который является наиболее из сложных задач в этой области. Развитие речевых технологий – это в первую очередь решение комплексов задач синтеза и распознавания разговорной речи, голосовое управление и формирование стиля речи, характерного в условиях отсутствия заранее заданной формы для удобного устного общения в диалоговом режиме.

Речь является естественным способом взаимодействия между людьми. В среднем человек способен произносить 150 слов в минуту, однако использует лишь около трети этого потенциала, говоря со скоростью 50 слов в минуту. Это свидетельствует о том, что человеческое общение посредством речи обладает большим потенциалом по скорости передачи информации, чем письменный формат. В связи с этим, возникла необходимость в создании технических средств, способных эффективно распознавать и интерпретировать речь человека, что привело к появлению систем вербальной коммуникации. Данные системы предназначены для оптимизации взаимодействия между человеком и машиной.

Вместо терминов «системы голосового управления» или «искусственный интеллект», в настоящей работе для обозначения указанных машин используется термин «системы устного диалога». Следует отметить, что с распространением подобных машин стал популярен ещё один общий термин – «автоматическое распознавание голоса и речи», подразумевающий преобразование речевого сигнала в текстовую форму.

Компании, работающие в сфере информационных технологий, понимают важность речи в разговоре и хотят улучшить качество общения со своими клиентами. Однако за всю 30-летнюю историю систем распознавания речи точность распознавания речи не была достаточно высокой, чтобы считать ее

надежной альтернативой человеческой речи. Поэтому использование систем голосового управления не стало объектом особого внимания для компаний.

Выбранная тема научного исследования, приобретая особую значимость в период ускоренного развития современных инновационных технологий, неоднократно становилась предметом обсуждения на международных, республиканских и вузовских конференциях, посвященных внедрению систем автоматического распознавания речи.

В связи с этим, рассмотрение процесса разработки, изучения и анализа автоматизированных программных систем распознавания речи с использованием передовых технологий программирования, подтверждает **актуальность** данной темы диссертационного исследования.

Степень изученности научной темы. В начале XX века, лишь в последние его десятилетия, некоторые разработки в области распознавания голоса и речи достигли приемлемого уровня точности, что привело к возобновлению интереса к этой области.

Развитие систем голосового управления можно разделить на три периода:

1. Период, основанный на правилах и шаблонах символов, который продолжался до конца 1990-х годов.

2. Период, основанный на статистическом обучении, охватывающий период с 1990-х по 2014 год.

3. Период, основанный на глубоком обучении, начавшийся в 2014 году и продолжающийся по настоящее время.

Что касается работы информационной системы и технологии обработки голоса и речи, для создания базы данных мы попытались изучить представляющие научную ценность материалы и идеи, предложенными отечественными и зарубежными учеными в этой области, использовать в своем исследовании.

Данная диссертационная тема получила значительное освещение в работах как отечественных, так и зарубежных ученых. Научные достижения этих исследователей в сфере обработки информационных систем послужили основополагающими для методологии настоящего исследования.

В числе наиболее важных источников следует выделить труды таких авторитетных ученых, как Дмитрий Богданов, Кенхен Чо, Йошуа Бенгио, Ян Хоровский, Дмитрий Сердюк, Филимон Бракель, Джанет М. Бейкер, Ли Дэн, Джеймс Гласс, Санджив Худанпур, Уильям Чан, Навдип Джейтли, Чжифэн Чен, Анджули Каннан, Ярин Гал, Зубин Гахрамани, Феликс Герс, Юрген Шмидхубер, Фред Камминс, Виталий Липчинский, Раджат Райна, Ананд Мадхаван, Эндрю Янг, Карл Уилмер Шольц, Джеймс С. Ирвин, Юнхуэй Ву, Майк Шустер, Чжифэн Чен, Alex Graves, Santiago Fernández, Faustino Gomez, Jürgen Schmidhuber и другие.

Разработка моделей, методов и алгоритмов распознавания речевых единиц и речи на основе нейронных сетей и машинного обучения имеет сравнительно недолгую историю. В своих работах ряд учёных, включая S.K. Gaikwad, B.W. Gawali, P.A. Yannawar, O. Abdel-Hamid, E.M. Johnson, A.A. Богустова, П.А. Парамонова, З.Д. Усманова, Исмоилова М.А. и Х.А. Худойбердиева, исследуют ключевые аспекты систем автоматического распознавания речи на естественном языке. Они рассматривают основные инструменты искусственного интеллекта, структуру голосового корпуса и технологии обработки практических примеров

машинного обучения, которые затем применяются в создании подобных систем на основе нейронных сетей.

Исследования российских учёных, таких как А.Б. Барский, А.А. Богустов, И.П. Меденников, С.В. Гребнов, И.Б. Тампель, А.А. Карпов и другие, посвящённые автоматической обработке единиц речи и распознаванию речи, дали точные и достоверные результаты. Работы этих учёных продемонстрировали возможности применения методов динамического программирования в анализе речи, а также позволили разработать систему распознавания голосовых команд, не зависящую от конкретного пользователя.

Так, в своей диссертации А.Б. Барский указывает на то, что появление глубокого обучения в области распознавания речи вызвало повышенный интерес со стороны коммерческих структур к внедрению подобных систем в собственной деятельности. Благодаря усовершенствованию методов глубокого обучения, системы распознавания речи достигли такого уровня точности, что стали применимы в реальных практических задачах и обладают коммерческим потенциалом [3].

Что касается исследований отечественных ученых и их научных достижений в области разработки информационных систем, которые легли в основу методологии исследования, могут увидеть на примере выпускников научной школы известного ученого в области информационных технологий, академика НАНТ – Усмонова З.Дж., ведущих научных сотрудников технологического университета Таджикистан – Исмоилова М.А., Зарипова С.А., Исмоиловой Р.М., Гуломсафдарова А.Г., научных сотрудников Таджикского технического университета имени академика М.С. Осими – Солиева О.М., Худойбердиева Х.А., Довудова Г.М., Косимова А.А., Ашурзода Б.Х. и других [4].

Один из наиболее значимых результатов в области распознавания ключевых слов в речи на таджикском языке был достигнут в работе Б.Х. Ашурзода, выполненной под руководством Х.А. Худойбердиева. В процессе исследования была определена система распознавания ключевых слов, основанная на невидимых моделях Маркова, и в данной исследовательской работе она также находит заданные слова, находящиеся в базе данных.

Следует отметить, что методы и приемы, упомянутые в данной диссертации, основаны на современной теории системы динамических объектов. В ходе диссертационного исследования автор опирался на основные принципы разработки систем компьютерного моделирования и нашел алгоритмическое решение задачи распознавания речи.

Данное исследование интересно тем, что оно основано на комплексном изучении платформ веб-приложений и существующих систем управления компьютерными программами и проведена их полная классификация. В последнее время распознавание речи привлекает все больше внимания как первый шаг в системе связи человека с компьютером, имеющий привлекательную коммерческую ценность.

С повышением потребностей в эффективной коммуникации с клиентами и ростом требований к внедрению систем голосового управления, многие компании, ориентируясь на опыт стран таких как США, Россия, Китай и др., провели тестирование систем голосовых команд. Согласно данным исследовательской

компания Gartner (компания, занимающаяся исследованиями в области информационных технологий), к 2021 году 15% всех взаимодействий с клиентами будут полностью автоматизированы с помощью искусственного интеллекта.

Дополнительное исследование показало, что 87% представителей бизнеса считают, что использование искусственного интеллекта в коммуникации приведёт к повышению удовлетворённости их бизнесом.

Несмотря на значительные успехи в области распознавания речи, получение его в необходимом объеме по-прежнему представляет собой сложную задачу.

Существует ряд факторов, затрудняющих полное понимание человеческой речи с помощью компьютерных технологий. Например, изменчивость речевого сигнала усложняет анализ непрерывной речи для машин.

Хотя наблюдается значительный прогресс в распознавании голоса и речи на таджикском языке, улучшения для других языков остаются актуальными. Успех, достигнутый в одном языке, не гарантирует аналогичного успеха в других.

В рамках данного исследования предпринята попытка разработать систему распознавания голоса и речи, которая может быть реализована в одной из областей деятельности финансовой компании или в сфере высшего профессионального образования.

Исследование показало, что значительная часть пользователей персональных компьютеров активно применяет передовые системы распознавания естественной речи, к числу которых относятся «Langsoft», «Lexical FreeNet», «LingSoft», «MonoConc/ParaConc», «Mystem», «netXtract», «Ngram Statistics Package – NSP», «Open Office Org», «Sakrament Text-to-Speech Engine», «Speech Synthesis and Recognition Laboratory», «Speech technology», «ASRaaS», «TTSaaS SaluteSpeech API», «Sputnik Voice», «Audiogram», «SmartNLP». Большое количество проектов используется в онлайн (на линии) и оффлайн (не на линии). Перечисленные системы разработаны и реализованы на основе методов нейронных сетей и машинного обучения. Такие проекты, как «SteosVoice», «Lovo.ai», «Zvukogram», «FreeTTS», «Robivox», «PlayHT», «Deepgram», «Murf.ai», расширяют рынок систем распознавания речи более чем на 40 естественных языках. Однако перечисленные проекты не могут поддерживать распознавание таджикской речи. Исходя из вышеизложенного, разработка проекта по распознаванию голосовых команд и речи на таджикском языке также требует использования нейронных сетей и методов машинного обучения.

Прежде всего, одним из главных преимуществ использования распознавания речи является сокращение человеческих ресурсов при обслуживании клиентов, что влечет за собой снижение финансовых затрат для сотрудников отрасли. Кроме того, персональный голосовой помощник может повысить доверие и удовлетворенность клиентов.

Исследования последних лет показывают, что глубокое обучение считается успешным в компании ASR. Среди существующих сетевых архитектур основное внимание в этом исследовании уделяется CNN и RNNs, а также обсуждаются важные компоненты систем глубокого обучения. Этими компонентами являются структура модели, большой обучающий набор, масштаб вычислений, предварительная обработка данных и настройка гиперпараметров. Также упоминается влияние этих компонентов на точность распознавания речи. В

частности, эта работа подкреплена детальным экспериментальным исследованием функциональных свойств нейронных сетей в задачах распознавания речи. В свою очередь, две разные модели нейронных сетей использовались для прогнозирования транскрипции речи на основе аудиоэкспериментов.

С глобальной точки зрения, конечная цель данного исследования — проложить путь к разработке масштабируемой и практичной модели распознавания голоса и речи, которая учитывает ограничения реального мира, модели, от которой могут извлечь выгоду как пользователи, так и компании.

В последние годы большинство исследователей используют различные методы и математические модели «распределения времени», основанные на теории государственных услуг по обработке информации в системе информационно-коммуникационных услуг.

В разработке компьютерно-математической модели информационной системы и сервисов таджикской базы данных распознавания речи были использованы принципы теории государственных услуг, функционирования телекоммуникационных сетей и основы математического моделирования. Адаптация модели к сетевой среде была осуществлена на основе исследований, проведенных в рамках проекта CNN-BLSTM европейскими учеными.

Связь исследования с программами (проектами) и научной тематикой. Разработка, анализ, методы и применение методологических основ и теории информационной системы и программных средств, представляющих собой точный алгоритм распознавания речи на таджикском языке и его реализация в достаточно быстрой системе для экспериментальной ориентации, в рамках реализации «Двадцатилетия изучения и развития естественных, точных и математических дисциплин в области науки и образования» на 2020-2040 годы была разработана программа «Электронное правительство». Кроме того, диссертационное исследование тесно связано с реализацией «Государственной программы по внедрению информационно-коммуникационных технологий в общеобразовательных учреждениях Республики Таджикистан» на 2018-2022 годы, концепцией формирования электронного правительства в Республики Таджикистан, Законами Республики Таджикистан «О государственном языке Республики Таджикистан», «Об информации», а также с осуществлением решения задач в планах-мероприятиях Государственной стратегии «Коммуникационно-информационных технологий для развития Республики Таджикистан», Стратегия инновационного развития Республики Таджикистан на период до 2020 года, Национальная стратегия развития Республики Таджикистан на период до 2030 года, что подтверждает актуальность научной темы [2].

ОБЩАЯ ХАРАКТЕРИСТИКА ИССЛЕДОВАНИЯ

Цель исследования – разработка математических и компьютерных моделей, методов обработки речи на основе алгоритмов машинного обучения и методов нейронных сетей, а также их реализация в компьютерной системе для автоматического распознавания голосовых команд на таджикском языке.

Задачи исследования. Для достижения поставленной цели исследования поставлены следующие задачи:

- изучение существующих научно-практических результатов в области технологий создания систем и комплексов распознавания голосовых команд;
- формулирование целей и задач исследования, разработка алгоритмических решений, математического и компьютерного моделирования процесса распознавания голосовых команд, оценка эффективности, тестирование и анализ полученных результатов;
- разработка и использование алгоритмов машинного обучения, методов нейронных сетей, математической статистики, компьютерного моделирования, вычислительной лингвистики и программного обеспечения для формирования цифрового портрета звуковых сигналов и их последующей обработки;
- обработка общедоступных математических и компьютерных моделей к базе данных таджикской архитектуры распознавания голоса и речи;
- разработка алгоритма подключения локальных гипертекстов к программной структуре модели распознавания речи;
- оценка качества разработанной таджикской системы распознавания голоса и речи на основе различных характеристик пользователей информационной системы.

Объект исследования - информационная система автоматизации и приложения баз данных для распознавания таджикской речи в форме архитектуры «клиент-сервер-клиент».

Предмет исследования - эффективные методы обработки, исследование, анализ и внедрение программно-алгоритмической системы и программных средств с распознаванием и автоматизацией распознавания таджикской речи.

Теоретические основы исследования. В рамках решения актуальных задач речевых технологий обработки речи, в частности, автоматическое распознавание речи на таджикском языке, применялись методы статистического анализа и скрытых марковских моделей, которые оказались не очень эффективными. Также в рамках исследования были проанализированы речевые корпуса русского и персидского языков с целью автоматического распознавания речи на таджикском языке, что показало уровень и надежность, значительно ниже требуемых. В исследовании использовались средства искусственного интеллекта, а именно современные методы машинного обучения и нейронные сети для высококачественного и быстрого распознавания команд в озвученной таджикской речи. Для достижения высокого уровня распознавания речи и его надежности впервые была разработана модель нейронной сети для машинного обучения с голосовыми единицами, произносимыми и записанными на таджикском языке. Для оценки эффективности процесса распознавания команд в озвученной таджикской речи использовались методы математической статистики и теории вероятностей.

Научная новизна исследования являются следующие пункты:

- на основе существующих акустических параметров речи на таджикском языке впервые представлена система распознавания голосовых команд, основанная на разработанную математических и компьютерных моделей;
- разработан и предложен двухэтапный алгоритм обучения акустических характеристик для повышения качества распознавания речи на основе глубоких нейронных сетей с учетом цифрового портрета голосовых команд в обучающем наборе;

- разработан речевая модель таджикского языка – TajNN-SAT на базе двухэтапного алгоритма обучения акустике с применением глубоких нейронных сетей;

- были разработаны логическая структура программы, рабочие процессы и доступные технологии распознавания голосовых команд на таджикском языке и реализованы в компьютерной модели;

- в первые смоделирован математическая модель рабочих процессов, на базе современных CASE-средств сформирована логическая структура системы и разработан программный комплекс Tajik-AVCR для распознавания голосовых команд на таджикском языке.

Положения, выносимые на защиту:

1. Математическая и компьютерная модели, методы выявления информационных параметров речи и её цифровой портрет с применением нейронных сетей.

2. Методы распознавания голосовых команд на таджикском языке, основанные на разработанную голосовую модель TajNN-SAT.

3. Разработан комплекс программ для распознавания команд на таджикском языке, которые эффективно используются в рабочих процессах в различных сферах общества, таких как образование, медицина и туризм.

Теоретическая и практическая значимость исследования. Теоретическая значимость данной диссертационной работы заключается в совершенствовании существующих алгоритмов и разработке новых алгоритмов обучения акустических паттернов на основе глубоких нейронных сетей для задач распознавания голоса и речи, а также в разработке и экспериментальном исследовании нового метода архитектуры информационных признаков, превосходящего используемые алгоритмы. Использование алгоритмических средств и разработка программного обеспечения при создании таджикской системы распознавания голоса и речи, показывают качество и скорость, достаточно высокие для использования в таких практических задачах, как, автоматическое озвучивание архивов фонограмм, поиск ключевых слов в речевом поток и группировку звуковых записей по темам. Основные результаты, полученные в диссертации, реализованы в программе «Система распознавания голоса таджикского языка», в учебном процессе ТГУПБП, программе DEM.tj при выполнении учебно-методической работы сотрудниками профессорско-преподавательского состава, преподавателей и студентов, а также в учебном процессе Института промышленности и сервиса.

Степень достоверности результатов диссертации. Основные результаты исследования получены в соответствии с его целями и задачами, основаны на методологии искусственного интеллекта. В этом процессе современные методы системного анализа, научная концепция компьютерной диагностики и компьютерного моделирования для машинного обучения нейронных сетей опираются на личный опыт автора и формируют надежную основу и содержание диссертации, выводы и рекомендации, изложенные в ней. Результаты научно-исследовательской работы были предоставлены на республиканском конкурсе «Наука – свет просвещение» в 2024-2025 годах в номинации «Искусственный интеллект и программирование» среди научных сотрудников, аспирантов и докторантов, преподавателей учебных заведений, инженеров, конструкторов и

представителей различных профессий, где было удостоено вторым местом. На основании полученных результатов было протестировано, усовершенствовано и внедрено программное обеспечение Tajik-AVCR. Основные результаты исследований используются в области науки и образования в Таджикском государственном университете права, бизнеса и политики и Институте промышленности и сервиса, в области медицины в ООО «Бахри Шифо» и ООО «Аль-Финур», в сфере туризма в ООО «Марвел Вита Травел» и туристической компании «Панжистон».

Соответствие диссертации паспорту научной специальности. Диссертация выполнена в соответствии со следующими пунктами Паспорта номенклатуры специальностей научных работников (технические науки): - по специальности 6D070400 – Вычислительная техника и программное обеспечение (6D070402 – Вычислительные машины, комплексы и компьютерные сети):

пункт 2 – «Теоретический анализ и экспериментальное исследование функционирования вычислительных машин, комплексов и компьютерных сетей с целью улучшения их технико-экономических и эксплуатационных характеристик» – компьютерная модель и комплекс компьютерных программ, разработанные в соответствии с данным пунктом, которые основаны на возможностях средств искусственного интеллекта и нейронных сетей, а также протестированы для надежной оценки методов и реализации комплексов;

пункт 3 – «Разработка научных методов и алгоритмов организации арифметической, логической, символьной и специальной обработки данных, хранения и ввода - вывода информации» – разработанная компьютерная модель, в соответствии с данным пунктом, которая позволяет организовать эффективную обработку голосовых данных и используется для автоматического распознавания команд в речевых единицах, произносимых на таджикском языке;

пункт 4 – «Разработка научных методов и алгоритмов организации параллельной и распределенной обработки информации, многопроцессорных, многомашинных и специальных вычислительных систем» – программный комплекс, разработанный на основе методов и алгоритмов обработки специальных цифровых данных для автоматического распознавания команд в таджикской речи, соответствует данному пункту;

Личный вклад соискателя учёной степени в подготовке диссертации выражается в самостоятельном написании диссертации, непосредственном участии в тематических дискуссиях, теоретических исследованиях, алгоритмизации, программировании, моделировании, проведении компьютерных экспериментов, анализе и обобщении выводов. Обработаны методы и алгоритмы распознавания голоса и речи, проведена эффективная экспериментальная оценка разработанных методов и алгоритмов. Разработаны программные средства, являющиеся ключевой частью таджикской системы распознавания голоса и речи. Основные научные положения и выводы диссертации были обоснованы и подтверждены с помощью метода компьютерного моделирования и программных модулей, разработанных в виде комплекса объектно-ориентированных программ.

Апробация и внедрение. Результаты исследования кандидатской диссертации обсуждались и публиковались на следующих научно-методических конференциях: Научно-практическая конференция ученых, профессоров и

молодых исследователей ТГУПБП (2019-2020-2021-2022 годы), научно-теоретическая конференция «Фундаментальные и прикладные науки»: состояние и развитие отрасли» (Петрозаводск, г. 2022 г.), научно-теоретическая конференция «Роль науки в становлении национальной государственности» (Худжанд, 2018 г.); «Развитие естественных и точных наук, математики в системе формирования цифровой экономики» (Худжанд, 2020 г.); научно-практическая конференция «Наука и инновации в системе реализации целей национальной стратегии» (Худжанд, 2021 г.); научно-методические конференции «Инновационные методы образования: современное состояние и перспективы совершенствования» (Худжанд, 2021 г.) и «Национальная концепция образования – основа воспитания творческого и инновационного поколения» (Худжанд, 2021 г.); научно-методическая конференция «Актуальные проблемы методики преподавания математических, точных и естественных предметов в ВУЗе: теория и практика»; (Худжанд, 2022 г.); научно-практическая конференция «Интеграция науки и производства для реализации целей национальной стратегии» (Худжанд, 2022 г.); научно-методическая конференция «Национальное единство – важный фактор развития государства и общества» (Худжанд, 2022 г.).

Публикации по теме диссертации. Результаты исследования отражены в 18 публикациях автора, из них 6 – научные статьи, опубликованные в рецензируемых журналах ВАК при Президенте Республики Таджикистан, остальные – статьи, опубликованы в других научных изданиях и материалах научных конференций. В частности, из Министерства культуры Республики Таджикистан получено Свидетельство о регистрации научных работ и авторских прав на Web-приложение «Система распознавания таджикских команд» от 23.02.2024 № 188, а также из «Национального патентно-информационного центра» при Министерстве экономического развития и торговли Республики Таджикистан получен Свидетельство о государственной регистрации информационных ресурсов на Web-приложение «Система распознавания голоса на таджикском языке» от 10.07.2024 № 4202400519.

Структура и объём диссертации. Диссертация состоит из разделов «Введение», «Общее описание работы», трех глав, «Заключение» и «Рекомендации», «Список использованной литературы» и «Публикации по теме диссертации соискателя учёной степени» и двух приложений. Общий объём работы состоит из 165 страниц компьютерного текста, набранного с помощью текстового процессора Microsoft Word, и включает 31 рисунков, 2 диаграммы и 16 таблиц. Список литературы включает 154 наименований.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении аргументированы актуальность темы исследования, описаны цели и задачи, предложены объект, предмет и методология исследования, отражены научная новизна, теоретическая и практическая значимость полученных результатов, а также приведены достоверность апробации и публикации полученные в диссертационной работе.

В **первой главе** «Теоретические основы моделей, методов и алгоритмов распознавания речи» рассмотрены вопросы, связанные с распознаванием

голосовых команд, с выявлением основных конструктивных параметров речевых элементов, а также сформированы задачи, которые необходимо решить при разработке систем обработки речи на таджикском языке.

В § 1.1 представлены общие сведения об актуальных моделях и методах распознавания речи, научно-практических результатах использования систем распознавания речи. Исследования показывают, что распознавание речи – это первая способность, с помощью которой системы искусственного интеллекта должны вести с человеком человеческий диалог. Распознаватель речи может распознавать человеческий голос и преобразовывать речь в текст, который представляет собой форму данных, читаемых компьютером. Это исследование является продолжением в области глубокого обучения распознаванию речи.

Научно-исследовательский результаты последних лет показывают, что системы распознавания речи внедряются в мобильных компаниях. Кроме того, её применение отчетливо видно в различных сферах образования, экономики, маркетинга, здравоохранения и туризма.

В § 1.2 представлена классификация моделей и методов распознавания речи в процессе распознавания путем выявления сходных признаков путем разделения изображений на общие классы. Суть этого метода в том, что изображения состоят из ряда свойств или общих характеристик и принадлежат к одному классу. Эти общие свойства можно импортировать в системную память.

Для решения задачи обработки звуковых сигналов и их преобразования в голос предоставлен алгоритм, который представлен на рисунке 1. В алгоритме при входе учитываются как, цифровой портрет элементов речи, так и акустическая модель выходного сегмента. Для каждого входного элемента создается статистическая модель, описывающий метрику и цифровой портрет его записи.

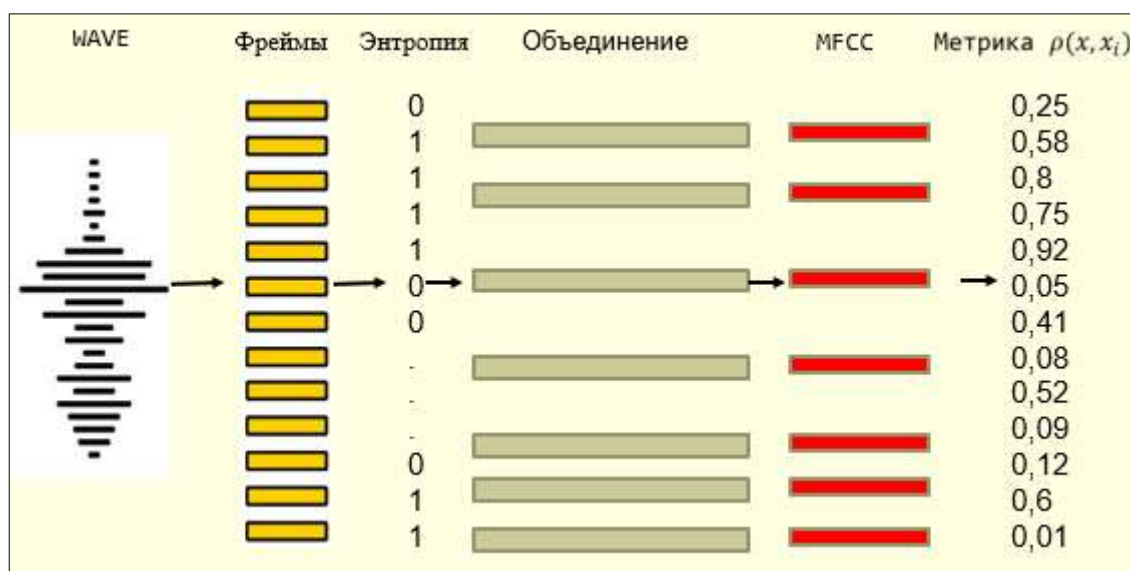


Рисунок 1 – Основной алгоритм цифрового распознавания речи

Система обработки предназначена для распознавания голосовых команд, произносимых отдельно. Распознавание слитной речи выходит за рамки исследования диссертации и требует передовых методов обработки. Система

распознавания голосовых команд представляет собой комплексный механизм, который состоит из алгоритма распознавания отдельных слов на базе методов динамического программирования. Такие требования, упомянутые выше, составляют содержание вопросов исследования.

Данная система обработки данных предназначена для идентификации индивидуальных голосовых команд. Анализ непрерывной речи выходит за рамки настоящего исследования и требует применения более сложных методов обработки. Система по своей сути является системой распознавания команд. В основе алгоритма лежит метод динамического программирования. В зависимости от типа структурной единицы, система должна быть способна идентифицировать отдельные слова.

Перечисленные выше требования определяют предмет данного исследования.

В § 1.3 представлены результаты практического использования алгоритма распознавания речи, основанного на динамическое изменение времени на примере речи таджикского языка. Вопрос включает в себя обработку голосовых единиц и автоматическое произношение речи на основе сопоставления с образцом. В памяти системы распознавания хранится множество изображений, принадлежащих одному классу. По мере появления новых изображений система сравнивает их одно за другим с изображениями, уже хранящимися в ее памяти (рис. 2).

Если в системе распознавания голосовых команд используется база данных звуковых элементов на примере отдельных букв, фонем, слогов, тогда её алгоритмический подход можно описать как основанный на последовательность коллекции речевых элементов. Такой подход эффективен в тех случаях, когда входные данные соответствуют заранее записанным звукам. Данный метод, является простым в реализации задачи, но в свою очередь ограничивает распознавания голосовых команд в узких областях, где входные данные полностью совпадают с эталонными. Метод перебора компонентов класса демонстрирует высокую точность при условии близости входных изображений к эталонным образцам.

В § 1.4 рассматриваются алгоритмы распознавания речи на основе нейронных сетей и их применение в вопросах использования алгоритмов распознавания речи на английском и русском языках. Также представлены результаты переработанных нескольких алгоритмов распознавания речи на таджикском языке. Нейросетевые системы распознавания речи используют сверточные нейронные сети, по крайней мере, на определенном этапе обработки. Относительно удачная структура цепных нейронных сетей для распознавания речи представляет собой длинную цепочку кратковременной памяти и считается одним из лучших методов описания моделей.

Для получения соответствующего результата используется функция распределения, учитывающая различные измерения, такие как пространство, время и даже семантика. Это также позволяет отслеживать интерес к входным данным, сравнивая их с заданным шаблоном на основе сходства или релевантности.

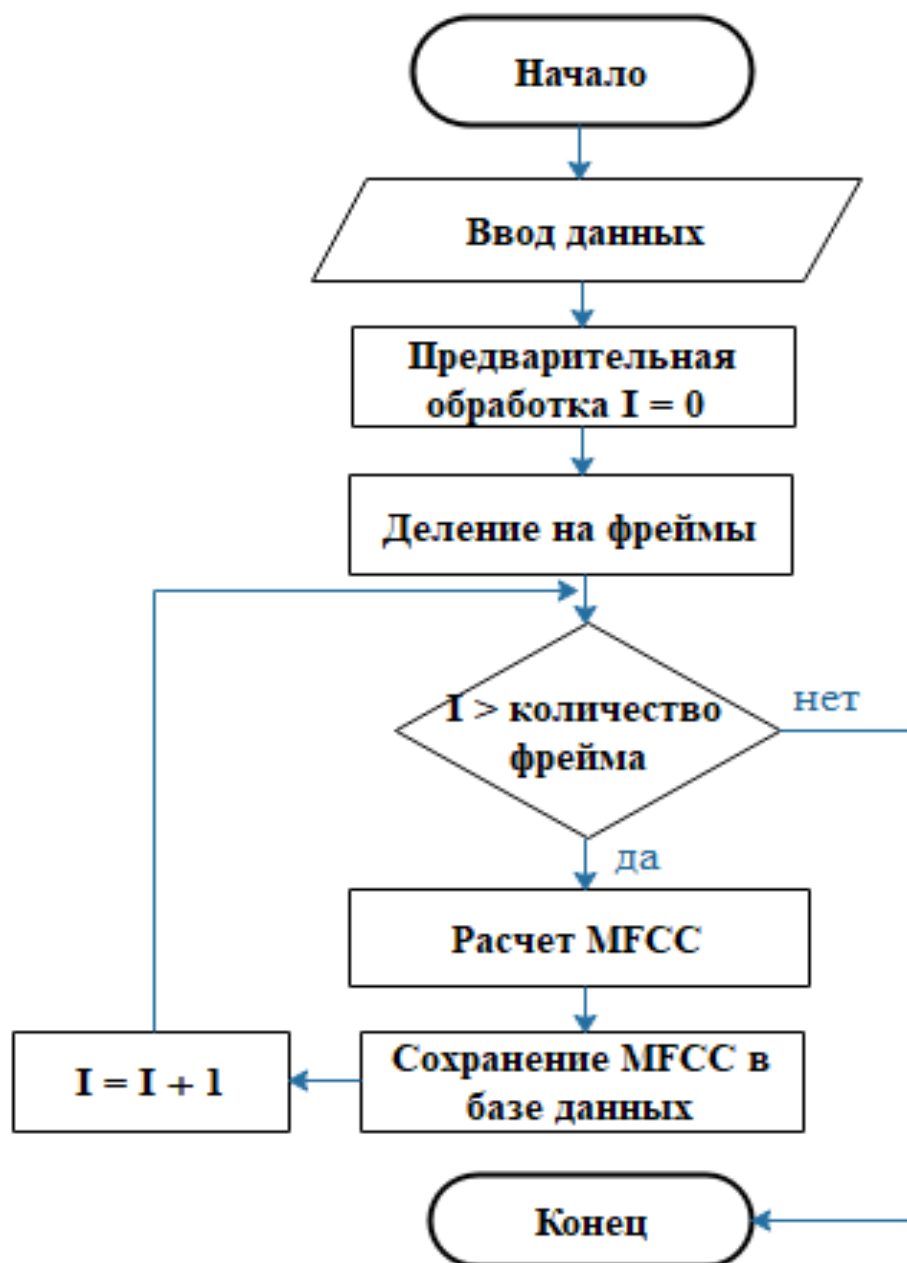


Рисунок 2 – Блок-схема ввода выборки в источник данных

На рисунке 3 показан пример машинного перевода, т.е. последовательность ввода информации на таджикском языке и экспорта ее на английский язык. Важно отметить, что внимание имеет два положительных эффекта: во-первых, оно улучшает производительность сети, а во-вторых, объясняет поведение нейронных сетей, которое трудно понять.

На рисунке 4 показаны три модуля модели TajNN-SAT, а ее пирамидальная структура изображена с двумя отдельными компонентами. Как видно на рисунке, декодер использует функцию принудительного обучения. Это означает, что выходной сигнал предыдущего слоя передается на текущий уровень декодирования в качестве входного.

Проигравший в модели TajNN-SAT может увидеть длинную пирамидальную цепочку кратковременного двустороннего кодирования входных данных памяти x

в функции высокого уровня h . В то же время декодер прослушивателя, основанный на внимании, создает символы y из h .

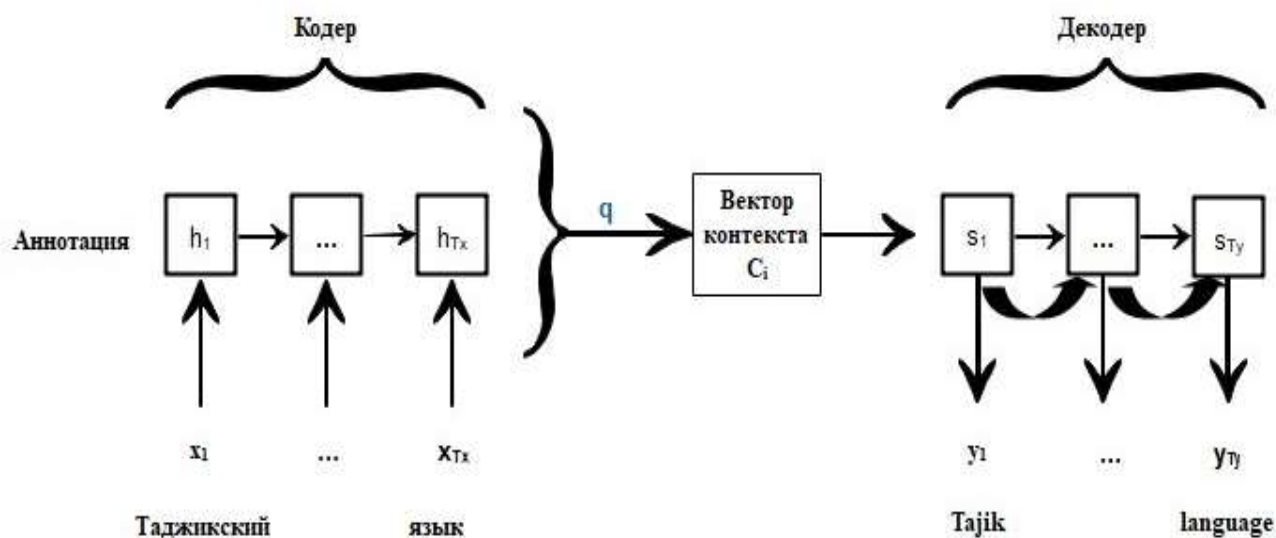


Рисунок 3. – Нейронная сеть кодер-декодер

Кодировщик использует длинную цепочку двусторонней кратковременной памяти пирамидальной структуры. Эта пирамидальная структура уменьшается с длиной h (представляющей самый высокий уровень входных данных x). Из-за сложности процесса декодирования экспорт соответствующих данных из большого количества временных шагов ввода очень медленно интегрируется длинной цепочкой простой двусторонней кратковременной памяти. Следовательно, полезно использовать длинноцепочечную пирамидальную форму двусторонней кратковременной памяти, поскольку это уменьшает размер времени декодирования и снижает вычислительную сложность. В результате модель интересов может получить соответствующую информацию за меньшее количество времени.

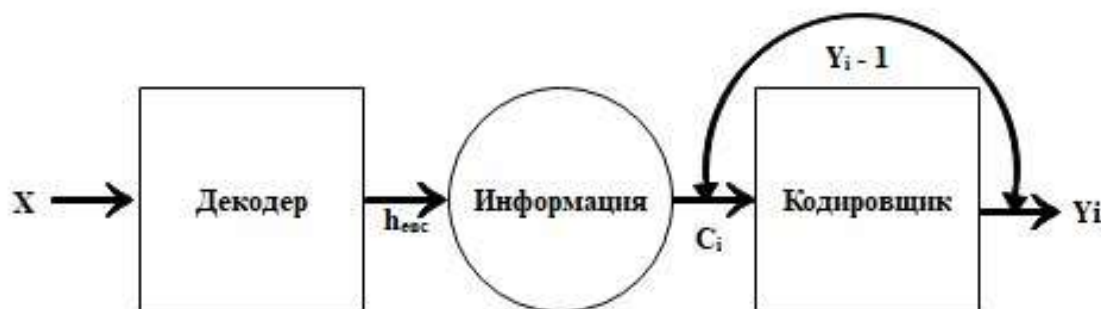


Рисунок 4. – Компоненты модели TajNN-SAT

Во второй главе «Проектирование моделей и методов распознавания речи» предложены результаты исследования, полученные в диссертационной работе, такие как математическая и компьютерная модель цифрового портрета голосовых команд на примере таджикского языка. На базе методов машинного обучения и

нейронных сетей рекомендованы способы распознавания речевого потока на таджикском языке с применением разработанной модели TajNN-SAT.

В § 2.1 приводится математическая модель цифрового речевого изображения, созданного для первичной обработки речи. Опыт показывает, что после первичной обработки речи набор кадров соответствует определенному слову. Для того чтобы сравнить сгенерированные данные и набор заданных слов, необходимо определить характеристики голоса. На основе заданных единиц вычисляется расстояние между векторами речевых кадров и векторами множества значений отражения задачи распознавания и определяется соответствие в зависимости от близости векторов.

В качестве метрики, например, можно принять среднеквадратичное значение, имеющее следующий вид:

$$\rho = \sqrt{\frac{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2}{n}} \quad (1)$$

где, n – количество принятых значений в наборе кадров, $x_i, y_i, i=(1,n)$ – компоненты векторов поля определения и отражения значений задачи распознавания, соответственно. Такая метрика не дает желаемого результата из-за большого значения n и отсутствия необходимой информации для процесса распознавания для любого значения $i, i=(1,n)$.

Человеческая речь – сложный сигнал, и эта трудность обусловлена наличием большого количества особенностей человеческой речи. Характеристики человеческой речи можно интерпретировать на разных уровнях, например акустическом, семантическом, лингвистическом, психологическом и т.д. У каждого человека уникальный голос, и даже когда один и тот же человек произносит одно и то же слово, результаты будут разными. Поэтому область компьютерного анализа речи является приоритетным направлением научных исследований.

В § 2.2 изложены методы анализа показателей речевых единиц, основанные на результатах исследования особенностей речевых единиц, произносимых на таджикском языке. Полученные результаты имеют разные амплитудные значения в зависимости от среды компьютерной программы для записи звука – организатора WAV-файла, в которой он был создан. Для унификации полученных данных, которые могут иметь как положительные, так и отрицательные значения, выполняется процесс нормализации данных. Для нормализации полученные данные можно разделить на наибольшее абсолютное значение амплитуд:

$$a_i = \frac{a_i}{\max(\text{abs}(\text{minVal}), \text{abs}(\text{maxVal}))}, \quad (2)$$

где,

a_i – значение амплитуды, взятое из файла, функция $\max()$ – это функция, которая возвращает наибольшее значение своих аргументов, функция $\text{abs}()$ – это функция, которая возвращает абсолютное значение своего аргумента, minVal и

maxVal – наименьшее и наибольшее значения амплитуды соответственно, которые берутся из файла. При этом все значения амплитуд берут значения из отрезка $[-1;1]$.

Первым шагом обработки после получения данных из файла является разделение его на фреймы (кадры). Фрейм (кадр) – это разделение голосовых данных на короткие временные интервалы. Разделение данных на фреймы осуществляется не последовательно, а друг над другом. То есть конец одного фрейма соответствует началу другого фрейма (рис. 5).

Причина разделения речи на кадры состоит в том, что выполнять сравнения на основе кадров гораздо проще, чем сравнения в каждый момент времени. Наложение первого и последнего кадров позволяет более точно настроить процесс анализа кадров.

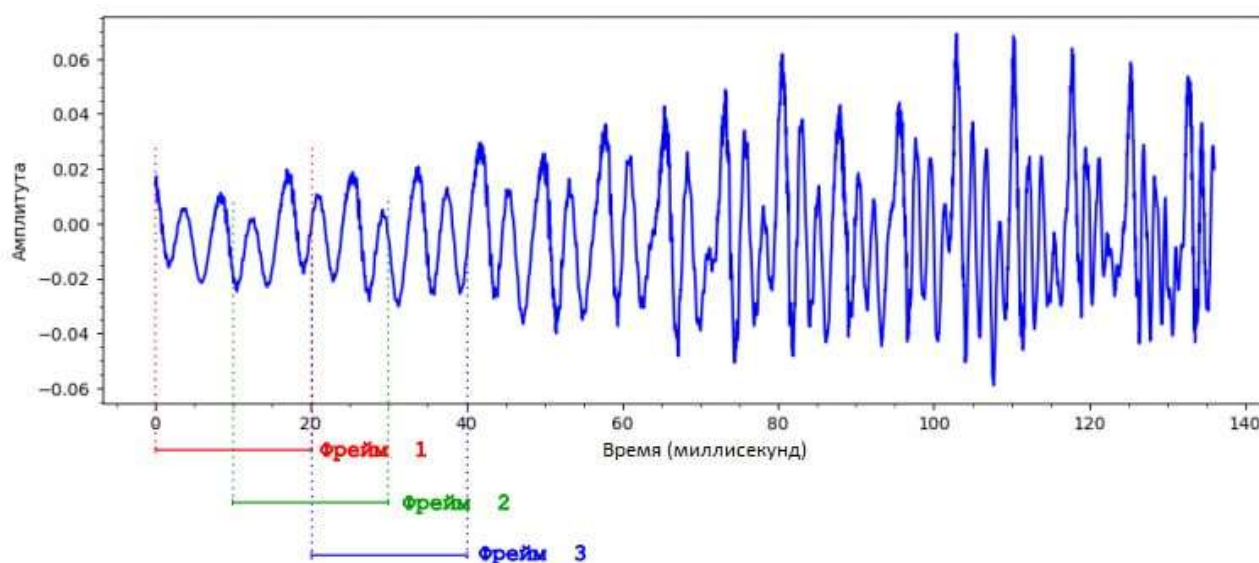


Рисунок 5 – Разделение голосовых данных на фреймы

В § 2.3 показаны методы структурирования и распознавания потока речи и их использование для решения основной задачи научной работы – распознавания команд на таджикском языке.

Как правило, глубокое обучение модели нейронной сети приводит к улучшению распознавания речи. Среди различных структур чаще всего используются глубокие обучения на основе алгоритма начальной цепи изучаемой акустики для нейронной сети и модели искусственного интеллекта, дающие значимые результаты на практике. В данном случае появляется задача формирования цепочки из акустической и языковой модели, а также цифрового портрета голосовых команд и текстовых элементов. Более того, комбинация слоев сверточной нейронной сети и длинной цепочки кратковременной памяти также хорошо работает для извлечения признаков. В этом подразделе диссертации представлена подробная информация о формах CNN-BLSTM и TajNN-SAT, распространенных в современной архитектуре между кодировщиками и декодерами.

В § 2.4 рассмотрены методы акустического обучения в модели TajNN-SAT, характеристики экспериментов и их результаты в сравнении с используемыми речевыми моделями. Сначала проводятся некоторые эксперименты для поиска

лучших функций и настройки их для модели TajNN-SAT. Для устранения временных ограничений и получения быстрых результатов, процесс анализа и эксперименты были проведены с набором данных Taj-Speech-Corpus.

По результатам исследовательской работы в диссертации был проведен комплекс заключительных экспериментов. Для повышения эффективности модели TajNN-SAT и CNN-BLSTM были изучены на наборах данных таджикской речи, то есть Taj-Speech-Corpus.

Таблица 1 - Характеристика модели CNN-BLSTM для набора данных Taj-Speech-Corpus

№	Максимальное соединение	Полноподключенный слой	Цепной уровень	Доступность данных	Частота ошибок кода доступа к данным %
1	Да	Да	1	Беспрерывный	0.81
2	Да	Да	1	Беспрерывный	0.81
3	Да	Да	5	Вероятный	0.78
4	Да	Да	5	Вероятный	0.80

Источник: Составлено автором.

Как видно из таблицы 1, процент частоты ошибок в кодах приходится на участие разработчиков. При этом слой длинной цепочки кратковременной двусторонней памяти имеет максимальный уровень интеграции и используется последнее соединение, которое и считается последним слоем. В слоях длинной цепочки двусторонней кратковременной памяти скрытый размер каждого направления равен 256.

Таблица 2 - Особенности модели CNN-BLSTM в наборах данных разных отраслей из Taj-Speech-Corpus

№	Максимальная конвергенция	Полноподключенный слой	Цепной уровень	Доступность данных	Массив данных	Ошибки в %
1	Да	Да	1	Беспрерывный	Медицина	0.85
2	Да	Да	1	Беспрерывный	Туризм	0.80
3	Да	Да	1	Беспрерывный	Образование	0.74
4	Да	Да	1	Беспрерывный	Детские сказки	0.78
5	Да	Да	5	Вероятный	Медицина	0.98
6	Да	Да	5	Вероятный	Туризм	0.80
7	Да	Да	5	Вероятный	Образование	0.76
8	Да	Да	5	Вероятный	Детские сказки	0.86

Источник: Составлено автором.

Как видно из таблицы 2, структурный результат CNN-BLSTM для набора данных Taj-Speech-Corpus не очень хорош по сравнению с результатом для набора данных медицинской отрасли. Это можно объяснить влиянием количества нулей, которые передаются в сеть при использовании медицинских слов в Taj-Speech-Corpus. Поскольку длина образцов в выборках Taj-Speech-Corpus совершенно разные, к более коротким выборкам добавляются нули, чтобы сделать входные данные одинаковой длины, чтобы все входные данные были одинаковой длины.

В целом используются не только базовые модели, в которых анализируются прогрессивные методы анализа и синтеза речи, но и некоторые модели, независимые на данный момент.

В третьей главе «Комплекс компьютерных программ TAJIK-AVCR» предоставлен процесс компьютерного моделирования обработки голосовых команд на таджикском языке с использованием объектно-ориентированного программирования. Предоставлены разработанные диаграммы с использованием унифицированного языка UML и компьютерная модель системы обработки текстовых данных и голосовых записей на таджикском языке с учетом вариантов использования, функций, структурных и физических объектов. А также обсуждены результаты экспериментов использования системы распознавания голосовых команд с оценкой её эффективности со стороны экспертов.

В § 3.1 представлена математическая модель рабочих процессов проекта распознавания речи на таджикском языке. Компьютерные программы TAJIK-AVCR представляют собой вычислительные проекты, позволяющие выделить непрерывную речь пользователя из общего потока звуков. Эти системы основаны на технологии распознавания речи, которая преобразует голос в цифровые сигналы посредством процесса распознавания слов. Также решением проблемы является разработка программных модулей распознавания речи и поиска необходимых данных в результирующем тексте [6].

Для начала необходимо установить структурную организацию системы. В рамках этой системы важными будут следующие элементы:

1. *Рабочие процессы* – *P*: регистрация и контроль пользователей; создание источника данных словаря; работа с записью голоса; распознавание речи; анализ результатов распознавания речи.

2. *Субъекты* – *S*: администратор; простой пользователь; лингвист (специалист); инженер.

3. *Функциональные возможности* – *F*: регистрация, вход, редактирование «рабочего стола» со стороны пользователей; работа с базой данных: ввод и редактирование; запись и первичная обработка голосовых команд; работа с голосовыми командами: диагностика и сравнение с текстовыми единицами в словаре; вычисление процентного соотношения представленной текстовой единицы с результатом распознавания речи; обеспечение информационной безопасности.

4. *База данных и ее таблицы* – *T*: пользователь информационной системы: порядковый номер пользователя, имя, фамилия, пол, дата рождения, адрес электронной почты, сведения об образовании, должность; текстовый элемент: порядковый номер текстового элемента, текст, тип (буква, слог, слово, словосочетание, предложение), дата ввода, предметная область, пояснения; элемент звука: порядковый номер звукового элемента, запись голоса, тип (wav, mp3), дата записи, порядковый номер пользователя, комментарии; распознавание речи: порядковый номер текстового элемента, порядковый номер голосового элемента, процент точного распознавания, статус распознавания, комментарии; результат анализа пользовательской активности: порядковый номер пользователя, количество записанных речевых элементов, средний процент распознавания речи.

Математическая модель описывает взаимосвязь между субъектами, функциями и данными в статистическом анализе:

$$I_{S_i}^{P_j} \Rightarrow f_k(t_1, t_2, t_3, \dots, t_q) \Rightarrow I_n \quad (3)$$

где:

$I_{S_i}^{P_j}$ – участие субъекта S(i) в процессе P(j) в рабочей среде I, $i=1,2,\dots,4$ и $j=1,2,\dots,5$;

$f_k(t_1, t_2, t_3, \dots, t_q)$ – наличие функции F(k) для обработки данных таблицы T(q), $k=1,2,\dots,14$ $q=1,2, \dots,5$.

I_n – представление результата в новой рабочей среде.

Математическая модель для реализации функции f1 может быть представлена следующим образом. Субъект S1 (*системный администратор*) запускает процесс P1 (*регистрация и управление пользователем*) и функцию f1 (*регистрация пользователя*), используя данные таблиц t1 (*пользователи информационной системы*). В результате генерируются новые данные I(n), которые используются для распознавания команд.

Пример математической модели с использованием представленных аргументов выглядит следующим образом:

$$I_{(S1)}^{(P1)} \Rightarrow f_1(t1) \Rightarrow I_n \quad (4)$$

где:

$I_{(S1)}^{(P1)}$ – интерфейс для регистрации пользователя;

$f_1(t1)$ – регистрация пользователя и сохранение данных в таблице пользователей информационной системы;

I_n – новые данные о записи и статусе звукового элемента.

В § 3.2 нашли свое решение вопросы логической структуры системы распознавания голосовых команд в комплексе Tajik-AVCR в компьютерном моделировании и предложения программного проекта за счет широкого использования CASE-инструментов. В этом направлении определяются и описываются с помощью специальных инструментов системные требования, логическая структура, диагностика объектов, логические связи между ними, рабочие процессы и фактическая модель [7].

Этап моделирования состоит из определения набора диаграмм с использованием языка UML для анализа спроса и предложения, которые важны для этапа обработки в реальном времени. Основные схемы, которые необходимы для проекта Tajik-AVCR, определены и представлены ниже (рис. 6).

В целом, разработанная система Tajik-AVCR имеет две особенности. Первая, сначала показываются участники, которые считаются его пользователями. Второй особенностью является идентификация всех вариантов использования в процессе распознавания голосовых команд, что определяет, с одной стороны, мониторинг

поведения информационной системы. Это основной этап процесса компьютерного моделирования информационных систем.

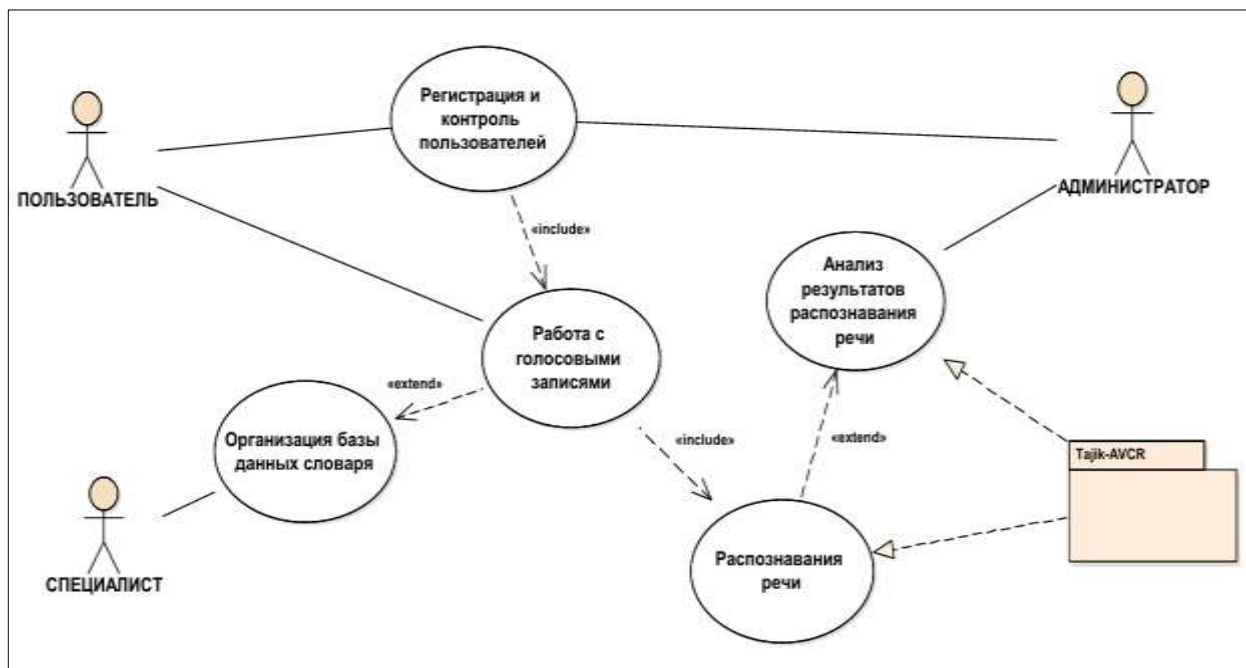


Рисунок 6 – Пример поведения системы Tajik-AVCR

В § 3.3 рассматривается практическое использование системы распознавания голоса Tajik-AVCR путем представления логической структуры системы распознавания голоса и пользовательского интерфейса (рис. 7).

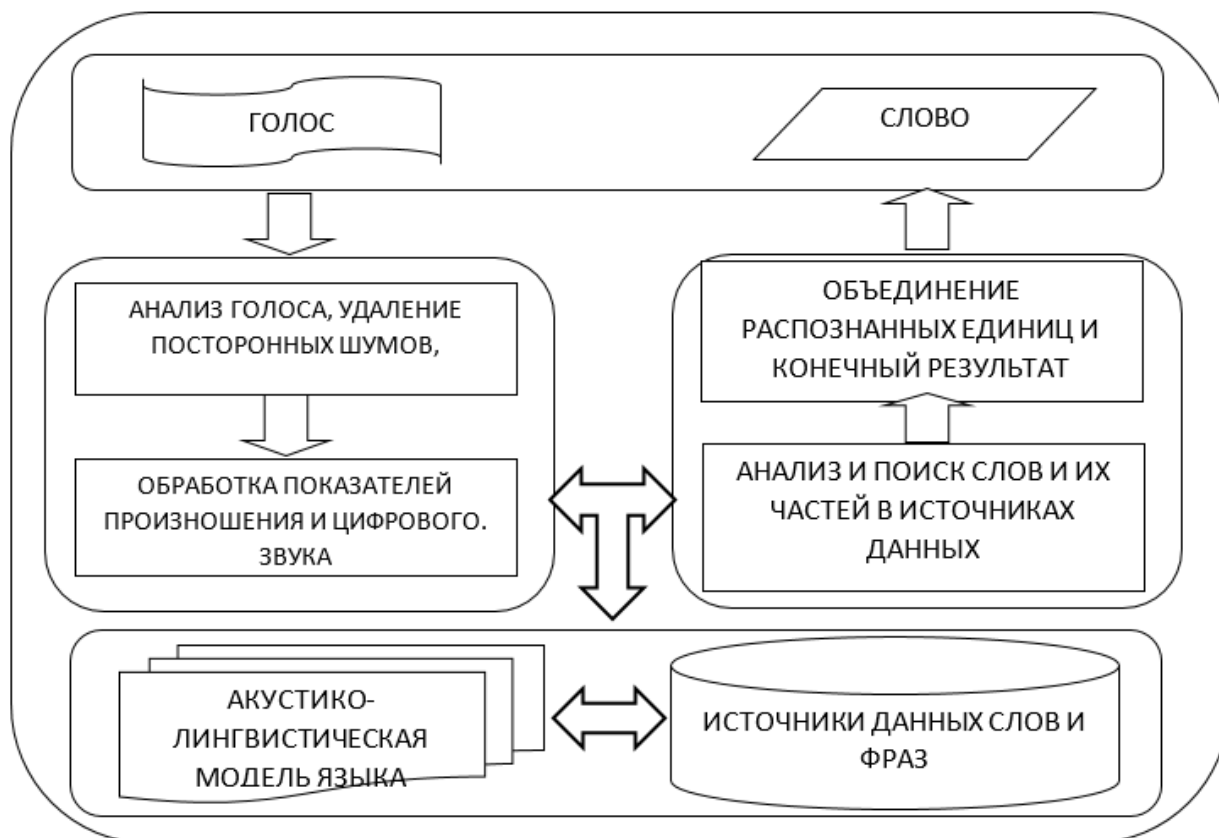


Рисунок 7 - Логическая структура системы распознавания голоса – Tajik-AVCR

При входе в программу принимается голосовая команда, что вы хотите выполнить поиск. В начале осуществляется отдельная обработка этих данных. Затем вводится задача обработки свойств произнесенной речи в виде цифрового звука. В случае обработки речи активируются акустическая и лингвистическая модели, подготовленные на основе модели таджикского языка.

После обработки входной речи программа отображает распознанный и найденный текст в специальной части текстового зеркала. Для активации всего потенциала системы распознавания корректируется положение программы. Наличие дополнительных программных модулей на компьютере пользователя является одним из его эффективных условий. Программа запускается с использованием рабочего файла Tajik-AVCR.exe. В главном интерфейсе доступны несколько блоков, в каждой из которых пользователь работает в соответствующей области: медицина, туризм, изучение языков, учебный процесс.

Интерфейс разработан для широкого круга пользователей и предоставляет им взаимодействия с программным комплексом, который разработан на базе визуального программирования. В данном случае пользовательский интерфейс представлен в упрощенном виде, что требует внесения некоторых изменений при реализации программного обеспечения (рис. 8).

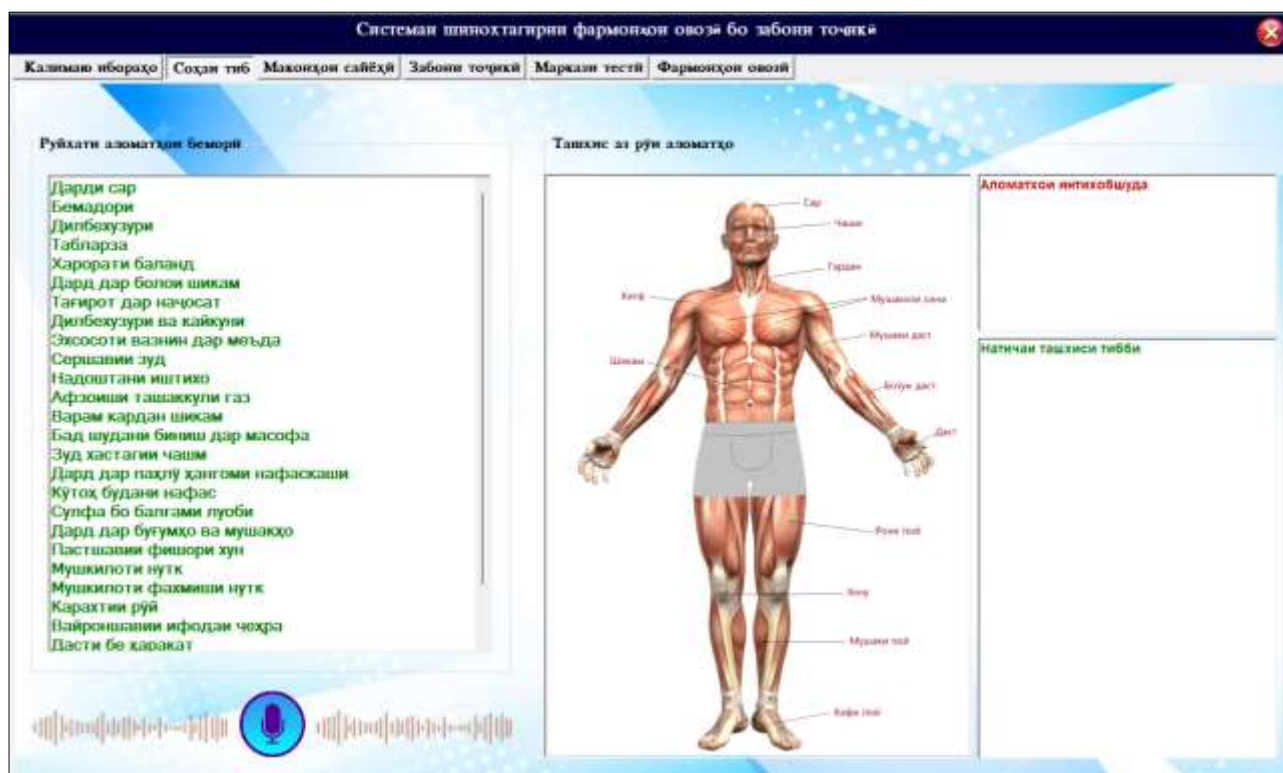


Рисунок 8 – Пользовательский интерфейс Tajik-AVCR в области медицины

В § 3.4 представлена оценка эффективности системы распознавания речи на основе результатов, полученных в практическом использовании в зависимости от ряда характеристик пользователя в комплексе программ Tajik-AVCR. Следует отметить, что средний процент распознавания речи среди тридцати пользователей составляет 76,48%. Диапазон процента распознавания речи составляет от 64,33 до 85,47 процента. Полученный результат, безусловно, отвечает требованиям

современных условий обработки систем автоматического распознавания речи и рынка программного обеспечения. Следует также отметить, что пользователи были разделены на две группы: мужчины и женщины. Также для конкретного анализа пользователи были разделены на три возрастные группы: от 14 до 24 лет, от 25 до 44 лет и от 45 до 60 лет.

Анализ результатов позволяет сделать вывод, что распознавание речи и представление текста средняя разница в правильном распознавании между мужчинами и женщинами составляет 2,36 процента. При этом возрастной анализ показывает, что тексты произносятся по-разному друг от друга, что соответствует средней точности распознавания 76,74%. На рис. 9 показан процент распознавания речи в зависимости от пола и возраста пользователей в виде гистограммы.



Рисунок 9 – Процент распознавания речи в системе Tajik-AVCR в зависимости от пола и возраста

При этом средний показатель точного распознавания речи выше у женщин в возрасте от 25 до 44 лет, а относительно низкий процент точности распознавания речи отмечается у мужчин в возрасте от 14 до 24 лет. Причина данного различия, на наш взгляд, связана с процессом распознавания речи и естественным стилем произношения пользователей.

На основании результатов исследования можно сделать вывод, что метод использования алгоритмов нейронных сетей для обнаружения голосовых команд в непрерывной речи считается осуществимым и эффективным. Также было установлено, что вероятность правильного обнаружения голосовой команды в произнесенном тексте зависит только от качества распознавания и наличия достаточной информации в тексте в определенной области. В целом, результаты показывают, что разработанная система распознавания голосовых команд эффективна и отвечает поставленным задачам в соответствии с современными требованиями.

ВЫВОДЫ

1. Основываясь на научно-практические исследования в направлении вычислительной лингвистики в Республике Таджикистан, а также за её рубежом **сформированы** как теоретические задачи, то есть разработка математических и компьютерных моделей, так и практических задач, то есть предоставление методов обработки голосовых команд на основе нейронных сетей и их применение в процессе автоматического распознавания речи на таджикском языке [1-М], [3-М], [7-М], [9-М].

2. В целях решения задачи обработки информации системы распознавания голосовых команд на таджикском языке **разработаны, переработаны и предложены** новые методы на основе нейронных сетей [2-М], [10-М], [12-М].

3. **Разработаны** математические и компьютерные модели цифрового образа речи, **созданы** методы обработки речевых единиц и **применены** в виде отдельных алгоритмов [2-М], [4-М], [17-М].

4. **Внедрены** распространенные методы, связанные с нейронными сетями, такие как CNN-BLSTM и TajNN-SAT, которые состоят из комплекса моделей и алгоритмов были **применены** к задачам распознавания речевых единиц на естественном языке [2-М], [4-М], [8-М], [11-М], [14-М], [15-М].

5. В диссертации **представлены** основные вопросы сбора, хранения, обработки и оцифровки голосовых единиц таджикского языка, а также их цифровой образ [5-М], [7-М], [12-М], [13-М], [17-М], [18-М].

6. В первые в Республике Таджикистан **разработан** программный комплекс TAJIK-AVCR для распознавания голосовых команд на таджикском языке, в котором **реализованы** результаты научного исследования, а также **доказаны** целесообразность использования разработанных математических и компьютерных моделей, методы нейронных сетей, алгоритмы обработки голосовых команд [4-М], [6-М], [16-М].

7. Научно-исследовательские результаты диссертационной работы по изучению особенностей таджикского языка **считаются** прочной основой для всех пользователей информационных технологий и представляют собой комплекс программ TAJIK-AVCR в области науки и образования в Таджикском государственном университете права, бизнеса и политики и Институте индустрии и услуг в области медицины в ООО «Бахри Шифа» и ООО «Аль-Финур», в сфере туризма в ООО «Марвел Вита Трэвел» и туристической компании «Панчистон», которые **обосновывают** практическое подтверждение результатов [2-М], [5-М], [15-М].

РЕКОМЕНДАЦИИ

Основными научными результатами диссертации являются дальнейшее развитие теории проектирования баз данных и решение научной проблемы, то есть разработка эффективных методов формирования информационных систем и приложений распознавания речи, которые используются при выполнении информационно-технических задач в учреждениях высшего профессионального образования, особенно в образовательном процессе, здравоохранении и туризме.

Практическая значимость диссертационного исследования заключается в разработке общего метода компьютерного моделирования баз данных

распознавания речи. Внесение определенных изменений в алгоритм и программную систему позволяет ввести проектирование моделей других образовательных учреждений в шаблон существующей модели. Полученная технология может внести ценный вклад в развитие практики компьютерного моделирования.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

[1]. Послание Президента Республики Таджикистан уважаемого Эмомали Рахмона об основных направлениях внутренней и внешней политики республики. 28.12.2024. www.president.tj.

[2]. Национальная стратегия развития Республики Таджикистан на период до 2030 года. Постановление Маджлиси намояндагон Маджлиси Оли Республики Таджикистан от 1 декабря 2016 года №636.

[3]. Барский, А.Б. Искусственный интеллект и интеллектуальные системы управления: монография / А.Б. Барский. - Москва: РУСАЙНС, 2022.- 185 с.

[4]. Косимов, А.А. Становление компьютерной лингвистики Таджикистана / А. А. Косимов. – Душанбе: Ирфон, 2021. – 102 с. – ISBN 978-99975-63-74-3. – EDN GBKWZO.

[5]. Белоруцкий, Р. Ю. Распознавание речи на основе сверточных нейронных сетей / Р.Ю. Белоруцкий, С.В. Житник. //Вопросы радиоэлектроники. – 2019. – №. 4. – С. 47-52.

[6]. Усманов, З. Д. Низомҳои худкори коркарди маълумот бо забони тоҷикӣ / З. Д. Усманов, Х. А. Худойбердиев. – Худжанд: Ирфон, 2022. – 186 с. – ISBN 978-99985—6624-8. – EDN ULPYYQ.

[7]. Худойбердиев, Х.А. Моделирование системы автоматической обработки текста на таджикском языке / Х. А. Худойбердиев // International Journal of Open Information Technologies. – 2023. – Т. 11, № 3. – С. 27-33. – EDN KRBOBH.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ СОИСКАТЕЛЯ УЧЁНОЙ СТЕПЕНИ

Статьи в журналах и научных изданиях, издаваемых ВАК при Президенте Республики Таджикистан

[1-М] Дадобаев, А.А. Математическая модель обработки голоса и свойств информации [Текст] / А.А. Дадобаев // Университетское письмо. Серия естественных и экономических наук. - Они движутся. - 2023 - № 1 (64) - С. 12-15. ISSN 2077-4974

[2-М] Дадобаев, А.А. Математическая модель распознавания голоса [Текст] / А.А. Дадобаев / Политехническое сообщение: Секция «Интеллект». Инновации. Инвестиции - 2023 - № (61) - С. 27-33. ISSN 2520-2235.

[3-М] Дадобаев, А.А. Алгоритм распознавания речи [Текст] / А.А. Дадобаев // Университетское письмо. Серия естественных и экономических наук. Худжанд. – 2024- № 1 (68) – С. 35–40. ISSN 2077–4974.

[4-М] Дадобаев, А.А. Программирование и разделение голосовых знаков речи [Текст] / А.А. Дадобаев / Политехническое сообщение: Секция «Интеллект». Инновации. Инвестиции - 2024 - № (64) - С. 53-57. ISSN 2520-2235.

[5-М] Дадобаев, А.А. Система поиска и распознавания голосовых команд в таджикской речи [Текст] / Х.А. Худойбердиев, А.А. Дадобаев, Ш.Н. Ашурова // Политехническое сообщение: Секция «Интеллект». Инновации. Инвестиции – 2024. – № 3 (58). - С. 181 – 189.

[6-М] **Анварзода А.А.** Особенности мобильных операционных систем и кроссплатформенных приложений для распознавания речи [Текст] / Б.Х. Ашурзода, А.С. Бобоев, **А.А. Анварзода** // Политехническое сообщение: Секция «Интеллект». Инновации. Инвестиции – 2025. – № 3 (71). – С. 63 – 67.

Статьи, опубликованные в других научных журналах

[7-М] **Дадобаев, А.А.** Методы распознавания речи и ее классификация [Текст] / Ш.М. Исмаилов, **А.А. Дадобаев** // Материалы международной научной конференции «Загадки и перспективы связи в современном мире», «Современная наука», ш. Петрозаводск, 2021. – С. 166-173.

[8-М] **Дадобаев, А.А.** Компьютерные программы, обеспечивающие распознавание и сегментацию речи [Текст] / **А.А. Дадобаев** // Материалы международной научной конференции «Науки», «Современная наука», ш. Петрозаводск, 2022. – С. 21 - 25.

[9-М] **Дадобаев, А.А.** Теоретическая классификация модели распознавания голоса [Текст] / **А.А. Дадобаев** // Материалы международной научной конференции «Лидеры СНГ», г. Ташкент, 2023. – С. 249 – 253.

[10-М] **Дадобаев, А.А.** Основные методы решения задач распознавания речи [Текст] / **А.А. Дадобаев** // Материалы международной научной конференции «Молодой исследователь СНГ», г. Ташкент, 2024. – С. 638 – 643.

[11-М] **Дадобаев, А.А.** Метод и алгоритм распознавания речи [Текст] / **А.А. Дадобаев** // Материалы научно-практической конференции профессоров, преподавателей и молодых исследователей Национальной академии наук «Наука и инновации для достижения национальных целей», г. Худжанд, «Дабир», 2019.-С. 111-114.

[12-М] **Дадобаев, А.А.** Классификация систем распознавания голоса [Текст] / **А.А. Дадобаев** // Материалы научно-практической конференции профессоров, преподавателей и молодых исследователей «Интеграция науки и производства для реализации национальных стратегических целей». - Худжанд: «Дабир», 2022.- С. 144-147.

[13-М] **Дадобаев, А.А.** Теоретическая классификация модели распознавания голоса [Текст] / Х.Ш. Гаюров, **А.А. Дадобаев** // Материалы республиканской научно-практической конференции «Миссия физической науки в развитии современной техники и оборудования». - Худжанд: «Дабир», 2023.- С. 113-117.

[14-М] **Дадобаев, А.А.** Классификация систем распознавания голоса [Текст] / **А.А. Дадобаев** // Материалы научно-практической конференции профессоров, преподавателей и молодых исследователей «Развитие науки и инноваций в условиях цифровизации».- Худжанд: «Дабир», 2023.- С. 153-158.

[15-М] **Анварзода А.А.** Методы и алгоритмы синтеза речи на основе текста. [Текст] / Худойбердиев Х.А., Ашурзода Б.Х., Анварзода А.А., Ашурова Ш.Н. // Вестник ПИТТУ имени академика М.С. Осими. – 2025. – № 3(35).

Свидетельства об авторстве и государственной регистрации информационных ресурсов

[16-М] **Дадобаев, А.А.** Веб-приложение «Система распознавания таджикских команд» [СОФТ] / **А.А. Дадобаев** // – 23.02.2024 – №188.

[17-М] **Дадобаев, А.А.** Веб-приложение «Система распознавания голоса на таджикском языке» [СОФТ] / **А.А. Дадобоев, Х.А. Худойбердиев, Д.Ф. Махкамов** // – 10.07.2024. – № 4202400519.

[18-М] **Дадобоев, А.А.** Web-приложение “Web-приложение taj-speech-corpus – Формирование речевых баз данных для таджикского языкового корпуса” [SOFT] / А.Т.Мақсуди, Х.А. Худойбердиев, **А.А. Дадобоев, Ш.Н. Ашурова, А.А. Назаров, Д.Ф. Махкамов** // – 17.09.2024. – № 4202400521.

АННОТАТСИЯ

ба автореферати рисолаи илмӣ Анварзода Акмал Анвар дар мавзӯи «Масъалаҳои шинохтагирии фармонҳои овозӣ ва татбиқи онҳо дар модели компютерӣ» барои дарёфти дараҷаи илмӣ доктори фалсафа (PhD) аз рӯи ихтисоси 6D070400 – Техникаи ҳисоббарор ва таъминоти барномавӣ (6D070402 – Мошинҳои ҳисоббарор, комплексҳо ва шабакаҳои компютерӣ)

Калимаҳои калидӣ: забони тоҷикӣ, амсиласозии математикӣ, технологияҳои овозӣ, забоншиносии компютерӣ, тарҳрезии компютерӣ, шинохти нутқ, фармонҳои овозӣ, зеҳни сунъӣ, омӯзиши мошинӣ, шабакаи нейронӣ, низомҳои иттилоотӣ, пояи додаҳо, технологияи барномарезӣ.

Мақсади таҳқиқот – таҳияи амсилаҳои математикӣ ва компютерӣ, усулҳои коркарди нутқ дар асоси шабакаҳои нейронӣ ва алгоритмҳое, ки барои шинохти худкори нутқи тоҷикӣ истифода карда шуда, ба роҳ мондани муколама бо забони табиӣ дар низомҳои “инсон-мошин” ва маҷмӯи барномаҳои компютерӣ.

Усулҳои таҳқиқот. Барои ҳалли вазифаҳои дар назди таҳқиқот гузошташуда усулҳои амсиласозии математикӣ ва компютерӣ, таҳлили системавӣ, таҳияи усулҳо, алгоритмҳо ва таъминоти барномавӣ барои шинохтагирии овоз ва нутқ, назарияи хизматрасонии оммавӣ, хулосабарории илмӣ, арзёбии экспертӣ, дастгоҳи аппарати математикӣ, забони барномасозии ба объект нигаронидашуда, технологияҳои инноватсионӣ, амнияти иттилоотӣ, озмоиши компютерӣ, таҳлили миқдорӣ ва сифатии натиҷаҳои таҳқиқот, умумигардонии натиҷаҳои таҳқиқот истифода шуданд.

Навгони илмӣ таҳқиқот. Дар натиҷаи кори илмӣ-таҳқиқотӣ бори нахуст дар заминаи пойгоҳҳои додаҳои мавҷуда бо баҳисобгирии хусусиятҳои маҳаллӣ барои шинохтагирии нутқи тоҷикӣ дар шакли як низоми мукаммали таҳқиқотӣ таҳия гардида, сохтори барномавӣ, тарзи кор ва технологияҳои дастрасӣ ба он мавриди таҳлил ва татбиқ қарор дода шудааст; барои сохтани хусусиятҳои иттилоотӣ, ки аз шабакаи амиқи нейронӣ бо гулӯи танг гирифта шудаанд, бо истифодаи мутобиқшавӣ ба гуфтор ва шароити акустикӣ фарқ мекунад, мавриди коркард қарор дода шудааст; алгоритми думарҳилаи ибтидоии акустикаи таълимӣ таҳия шудааст, ки дар асоси шабакаҳои амиқи нейронӣ, бо назардошти шумораи мисолҳои муқаррарӣ дар маҷмӯи омӯзиш фарқ мекунад ва ҳамзамон барои баланд бардоштани сифати шинохтагирии нутқ мусоидат мекунад.

Муқаррароте, ки барои дифоъ пешниҳод карда мешавад:

1. Амсилаҳои математикӣ ва компютерӣ, усули сохтани хусусиятҳои иттилоотии нутқ ва симои рақамии он дар асоси шабакаи амиқи нейронӣ муайян карда шуданд.

2. Алгоритми думарҳила барои оғоз кардани омӯзиши амсилаҳои акустикӣ дар асоси шабакаҳои амиқи нейронӣ ташаккул ва аз нав коркард карда шуданд.

3. Усулҳои сохторбандии низомии шинохтагирии нутқи забони тоҷикӣ ва омӯзиши мошинӣ дар амсилаи ЗИАО-ШН ва TajNN-SAT ташкил ва пешниҳод карда мешавад.

4. Бори аввал барои ҳалли масъалаи шинохти нутқи тоҷикӣ дар асоси шабакаҳои нейронӣ дар маҷмӯи барномаҳои компютерии TAJIK-AVCR бо пешниҳоди амсилаи математикии равандҳои корӣ ва сохтори мантикии низоми шинохтагирии нутқ дар асоси CASE-воситаҳо таҳия карда шуданд.

5. Натиҷаҳои истифодаи амалии низоми коркардшудаи шинохтагирии фармонҳо бо забони тоҷикӣ бо мақсади баҳодиҳии самаранокии онҳо дар соҳаҳои гуногуни ҷамъият ба монанди таълим, сайёҳӣ, тандурустӣ озмоиши таҷрибавӣ гузаронида шуданд.

АННОТАЦИЯ

к автореферату научной диссертации Анварзода Акмала Анвара на тему «Проблемы распознавания голосовых команд и их применение в компьютерной модели» на соискание ученой степени доктора философии (PhD) по специальности 6D070400 – Вычислительная техника и программное обеспечение (6D070402 – Вычислительные машины, комплексы и компьютерные сети).

Ключевые слова: таджикский язык, математическое моделирование, голосовые технологии, компьютерная лингвистика, компьютерный дизайн, распознавание речи, голосовые команды, искусственный интеллект, машинное обучение, нейронная сеть, информационные системы, база данных, технология программирования.

Цель исследования – разработка математических и компьютерных моделей, методов обработки речи на основе нейронных сетей и алгоритмов, используемых для автоматического распознавания таджикской речи, установления диалога с естественным языком в системах «человек-машина» и комплексе компьютерных программ.

Методы исследования. Для решения исследовательских задач используются методы математического и компьютерного моделирования, системный анализ, разработка методов, алгоритмов и программного обеспечения для распознавания голоса и речи, теория государственных услуг, научные выводы, экспертная оценка, математический аппарат, объектно-ориентированный язык программирования, инновационные технологии, информационная безопасность, компьютерное тестирование, количественный и качественный анализ результатов исследований, обобщение результатов исследований.

Научная новизна исследования. В результате научно-исследовательской работы впервые на основе существующих баз данных с учетом местных особенностей по распознаванию таджикской речи была разработана полноценная исследовательская система, определены структура программы, методы работы и технологии доступа. анализируется и реализуется; был обработан для создания информационных функций, полученных из глубокой нейронной сети с узким каналом, с использованием адаптации к речи и различным акустическим условиям; Алгоритм двухэтапного обучения акустике разработан на основе глубоких нейронных сетей с учетом количества регулярных примеров в обучающей выборке и при этом способствует повышению качества распознавания речи.

Положения, выносимые на защиту:

1. Определены математическая и компьютерная модели, метод создания информационных признаков речи и ее цифрового изображения на основе глубокой нейронной сети.
2. Разработан и переработан двухэтапный алгоритм запуска обучения акустических моделей на основе глубоких нейронных сетей.
3. Организованы и представлены методы военного структурирования распознавания речи на таджикском языке и машинного обучения, такие как CNN-BLSTM и TajNN-SAT.
4. Впервые для решения задачи распознавания таджикской речи на основе нейронных сетей разработаны компьютерные программы TAJIK-AVCR с представлением математической модели рабочих процессов и логической структуры системы распознавания речи на базе CASE-средств.
5. Проведены экспериментальные проверки результатов практического использования разработанной системы распознавания команд на таджикском языке с целью оценки ее эффективности в различных сферах жизни общества, таких как образование, туризм, здравоохранение.

ANNOTATION

**to the abstract of the Anvarzoda Akmal Anvar's scientific dissertation
on the topic "Problems of recognition of voice commands and their application
in a computer model" for the degree Doctor of Philosophy (PhD) in the specialty
6D070400 – Computing and software (6D070402 – Computing Machines,
Complexes, and Computer Networks)**

Keywords: Tajik language, mathematical modeling, voice technologies, computer linguistics, computer design, speech recognition, voice commands, artificial intelligence, machine learning, neural network, information systems, database, programming technology.

The aim of the research is to develop mathematical and computer models, speech processing methods based on neural networks and algorithms used for automatic recognition of Tajik speech, establishing a dialogue with natural language in human-machine systems and a complex of computer programs.

Research methods. In order to solve research problems, the following methods are used: mathematical and computer modeling, system analysis, development of methods, algorithms and software for voice and speech recognition, theory of public services, scientific conclusions, expert assessment, mathematical apparatus, object-oriented programming language, innovative technologies, information security, computer testing, quantitative and qualitative analysis of research results, generalization of research results.

Scientific novelty of the research. The result of the research work, for the first time, based on existing databases, taking into account local features for recognition of Tajik speech, a full-fledged research system was developed, determined the program structure, operating methods and access technologies analyzed and implemented; was processed to create information functions obtained from a deep neural network with a narrow channel, using adaptation to speech and various acoustic conditions; The two-stage acoustics learning algorithm is developed based on deep neural networks, taking into account the number of regular examples in the training sample and at the same time contributes to improving the quality of speech recognition.

The following results submitted for defense:

1. The mathematical and computer models, the method of creating information features of speech and its digital image based on a deep neural network are defined.
2. A two-stage algorithm for launching training of acoustic models based on deep neural networks is developed and revised.
3. Methods of military structuring of speech recognition in the Tajik language and machine learning, such as CNN-BLSTM and TajNN-SAT are organized and presented.
4. For the first time, TAJIK-AVCR computer programs with a presentation of a mathematical model of work processes and a logical structure of a speech recognition system based on CASE tools are developed to solve the problem of Tajik speech recognition based on neural networks.
5. Experimental tests of the results of practical use of the developed command recognition system in the Tajik language are carried out in order to assess its effectiveness in various areas of society, such as education, tourism, and healthcare.