

ХЎЖАЕВ ОТАБЕК КАДАМБАЕВИЧ

**ҚАРОР ҚАБУЛ ҚИЛИШГА КЎМАКЛАШИШ ТИЗИМЛАРИ УЧУН
ИНТЕЛЛЕКТУАЛ ТАҲЛИЛ ДАСТУРИЙ ВОСИТАСИНИ ИШЛАБ
ЧИҚИШ**

05.01.04-Ҳисоблаш машиналари, мажмуалари ва компьютер тармоқларининг математик ва
дастурий таъминоти

**ТЕХНИКА ФАНЛАР БЎЙИЧА ФАЛСАФА ДОКТОРИ (PhD)
ДИССЕРТАЦИЯСИ АВТОРЕФЕРАТИ**

ХЎЖАЕВ ОТАБЕК КАДАМБАЕВИЧ

**ҚАРОР ҚАБУЛ ҚИЛИШГА КЎМАКЛАШИШ ТИЗИМЛАРИ УЧУН
ИНТЕЛЛЕКТУАЛ ТАҲЛИЛ ДАСТУРИЙ ВОСИТАСИНИ ИШЛАБ
ЧИҚИШ**

05.01.04-Ҳисоблаш машиналари, мажмуалари ва компьютер тармоқларининг математик ва
дастурий таъминоти

**ТЕХНИКА ФАНЛАР БЎЙИЧА ФАЛСАФА ДОКТОРИ (PhD)
ДИССЕРТАЦИЯСИ АВТОРЕФЕРАТИ**

Техника фаилари бўйича фалсафа доктори (PhD) диссертацияси мавзуси Ўзбекистон Республикаси Вазирлар Маҳкамаси ҳузуридаги Олий аттестация комиссиясида В2017.4.PhD/T53 рақам билан рўйхатга олинган.

Диссертация Тошкент ахборот технологиялари университетидида бажарилган.

Диссертация автореферати уч тилда (Ўзбек, рус, инглиз (резюме)) Илмий кенгаш веб-саҳифасида (www.tuit.uz) ва «Ziyouet» Ахборот таълим порталида (www.ziyouet.uz) жойлаштирилган.

Илмий раҳбар: **Нишанов Аҳрам Хасанович**
техника фаилари доктори

Расмий оппонентлар: **Мўминов Баҳодир Болтаевич**
техника фаилари доктори

Хантов Файзулла Норбўтаевич
техника фаилари номзоди, доцент

Етакчи ташкилот: **Тошкент давлат техника университети**

Диссертация химояси Тошкент ахборот технологиялари университети ҳузуридаги DSc.27.06.2017.T.07.01 Илмий кенгашининг 2019 йил «14» июн соат 14⁰⁰ даги мажлисида бўлиб ўтади. (Манзил: 100202, Тошкент шаҳри, Амир Темур кўчаси, 108-уй. Тел.: (99871) 238-64-43, факс: (99871) 238-65-52, e-mail: tuit@tuit.uz).

Диссертация билан Тошкент ахборот технологиялари университети Ахборот-ресурс марказида танишиш мумкин (~~150~~ рақам билан рўйхатга олинган). (Манзил: 100202, Тошкент шаҳри, Амир Темур кўчаси, 108-уй. Тел.: (99871) 238-65-44).

Диссертация автореферати 2019 йил «01» июн да таркатилди.
(2019 йил «02» май даги 9 рақамли реестр баённомаси.)



Р.Х.Хамдамов
Илмий даражалар берувчи илмий кенгаш раиси, т.ф.д., профессор

Ф.М.Нуралиев
Илмий даражалар берувчи илмий кенгаш илмий котиби, т.ф.д., доцент

Х.Н.Зайниддинов
Илмий даражалар берувчи илмий кенгаш қошидаги илмий семинар раиси, т.ф.д., профессор

Диссертация мавзусининг долзарблиги ва зарурати. Жаҳонда дастурий маҳсулот ишлаб чиқариш саноатининг ривожига келажакда ахборот ва қарор қабул қилишга кўмаклашиш тизимларига алоҳида эътибор қаратмоқда. 2016 йилда биргина тиббиёт соҳасида беморларнинг электрон карталарини ишлаб чиқиш билан боғлиқ бўлган дастурий таъминотлар бозори айланмаси 20,55 миллиард долларни ташкил этган. «Ҳозирда «Open EMR» очик кодли тиббиёт ахборот тизимидан 10 дан ортик давлат, 1500 дан ортик шифохоналар амалда фойдаланадилар»¹. Жамиятдаги турли соҳаларнинг ривожланишига йўналтирилган ахборот ва қарор қабул қилишга кўмаклашувчи дастурий таъминотлар, уларга хизмат қилувчи махсус дастурий воситаларни ишлаб чиқишга АҚШ, Германия, Франция, Буюк Британия, Япония, Австралия, Жанубий Корея, Хитой, Ҳиндистон, Россия Федерацияси ҳамда Ўзбекистонда жуда катта эътибор қаратилмоқда.

Жаҳонда ахборот ва қарор қабул қилишга кўмаклашиш тизимларини қўллаш, яъни ахборот тизимлари маълумотлар базасидаги катта массивли маълумотларни интеллектуал таҳлил қилиб, уларнинг белгилари орасидаги боғлиқликларни аниқлаш, ўқшаш объектларни саралаш, вақтли қаторларда башоратлаш каби масалаларни ечиш орқали тегишли соҳа фаолиятидаги иш самарадорлигини оширишга доир кўплаб илмий изланишлар олиб борилмоқда. Бу борада жумладан, маълумотлар базасини бошқариш тизимларидаги маълумотларни интеллектуал таҳлил қилиш усуллари ва алгоритмларидан фойдаланган ҳолда қайта ишлаб қарор қабул қилишга кўмаклашувчи дастурий воситаларни ишлаб чиқиш муҳим вазифалардан бири ҳисобланади.

Республикамызда ахборот коммуникация технологиялари ва компьютерли диагностика воситалари асосида тиббий хизмат кўрсатишни яхшилаш, хусусан, кардиология соҳасидаги хизматлар сифатини оширишга алоҳида эътибор қаратилмоқда. 2017-2021 йилларда Ўзбекистон Республикасини янада ривожлантириш бўйича Ҳаракатлар стратегиясида, жумладан, «...соғлиқни сақлаш соҳасини, энг аввало, унинг аҳолига тиббий ва ижтимоий-тиббий хизмат кўрсатиш қулайлиги ҳамда сифатини оширишга қаратилган дастлабки бўғинини, тез ва шошилич тиббий ёрдам тизимини янада ислоҳ қилиш, аҳоли ўртасида соғлом турмуш тарзини шакллантириш, тиббиёт муассасаларининг моддий-техника базасини мустаҳкамлаш, ... ижтимоий соҳа, бошқарув тизимига ахборот-коммуникация технологияларини жорий этиш»² вазифалари белгиланган. Мазкур вазифаларни амалга ошириш, хусусан, кардиология соҳасидаги ахборот тизимларига интеллектуал таҳлил дастурий модулларини қўллаган ҳолда шифокорларнинг иш фаолияти самарадорлигини янада ошириш энг муҳим

¹ http://open-emr.org/wiki/index.php/OpenEMR_Professional_Support

² Ўзбекистон Республикаси Президентининг 2017 йил 7 февралдаги ПФ-4947-сон «Ўзбекистон Республикасини янада ривожлантириш бўйича Ҳаракатлар стратегияси тўғрисида»ги Фармони

масалалардан биридир.

Ўзбекистон Республикаси Президентининг 2017 йил 7 февралдаги ПФ-4947-сон «Ўзбекистон Республикасини янада ривожлантириш бўйича Ҳаракатлар стратегияси тўғрисида»ги Фармони, 2018 йил 2 августдаги ПК-3894-сон «Ўзбекистон Республикасида соғлиқни сақлашни бошқаришнинг инновацион моделини жорий этиш чора-тадбирлари тўғрисида»ги, 2012 йил 21 мартдаги ПК-1730-сон «Замонавий ахборот-коммуникация технологияларини жорий этиш ва янада ривожлантиришнинг чора-тадбирлари тўғрисида»ги қарорлари ва Вазирлар Маҳкамасининг 2012 йил 1 февралдаги 24-сонли «Жойларда компьютерлаштириш ва ахборот-коммуникация технологияларини бундан кейинги ривожлантиришга шароитлар яратиш учун чора-тадбирлар тўғрисида»ги қарори ҳамда мазкур фаолиятга тегишли бошқа меъёрий-ҳуқуқий ҳужжатларда белгиланган вазифаларни амалга оширишда ушбу тадқиқот муайян даражада хизмат қилади.

Тадқиқотнинг Республика фан ва технологиялари ривожланишининг устувор йўналишларига мослиги. Мазкур тадқиқот Республика фан ва технологиялари ривожланишининг IV. «Ахборотлаштириш ва ахборот-коммуникация технологияларини ривожлантириш» устувор йўналиши доирасида бажарилган.

Муаммонинг ўрганилганлик даражаси. Қарор қабул қилишга кўмаклашиш тизимлари, ахборот тизимлари, маълумотларнинг интеллектуал таҳлили, тимсолларни аниқлаш, сервисга йўналтирилган дастурий воситаларни лойиҳалаш ва ишлаб чиқиш, маълумотларга қайта ишлов бериш усуллари ва математик моделлаштириш масалаларининг ривожланиши ва уларнинг қўлланилиши бўйича ихтисослаштирилган илмий муассасаларда кўплаб олимлар илмий-амалий тадқиқотлар олиб боришган.

Жумладан, С.С. Aggarval, Е. S. Berner ишларидаги ахборот ва қарор қабул қилишга кўмаклашиш тизимлари ва уларни қуриш моделлари, С. М. Bishop, R. O. Duda, P. E. Hart, В. Н. Вапник, В. И. Васильев, Ю. И. Журавлев, А. А. Барсегян, Н. Г. Загоруйко, Г. С. Лбов, Л. А. Растрингин ишларида тимсолларни таниб олиш, маълумотларнинг интеллектуал таҳлилида синфларга ажратиш, вақтли қаторларни башоратлаш, объектларни синфлаштириш масалалари ўрганиб чиқилди.

Шунингдек, Республикамизда М. М. Камиллов, Т. Ф. Бекмуратов, Ш. Х. Фозилов, Р. Х. Хамдамов, М. А. Рахматуллаев, Н. А. Игнатьев, Д. Т. Мухамедиева, А. Х. Нишанов, Р. Н. Усманов, А. Р. Ахатов, А. И. Назаров каби олим ва тадқиқотчилар бу йўналишдаги илмий тадқиқотларга ўз хиссаларини қўшишган.

Ҳозирги кунгача ахборот тизимлари маълумотларини интеллектуал таҳлил қилиш мақсадида сервисга йўналтирилган архитектура асосида дастурий воситаларни ишлаб чиқиш ва турли платформалар учун ечимларни тақдим қилиш масалалари етарли даражада ўрганилмаган.

Диссертация тадқиқотининг диссертация бажарилган олий таълим муассасасининг илмий-тадқиқот ишлари режалари билан боғлиқлиги. Диссертация тадқиқоти Муҳаммад ал-Хоразмий номидаги Тошкент ахборот технологиялари университетининг илмий-тадқиқот ишлари режасининг ФА-А17-Ф056 «Объект ҳақидаги ахборот ноаниқ шароитида маълумотларни интеллектуал таҳлил қилиш учун мослашувчан тимсолларни таниб олиш тизимининг алгоритмик таъминоти» (2009-2012); БА-5-017 «Маълумотларни интеллектуал таҳлил қилиш усуллари асосида кўкрак бези ва бачадон бўйни саратонини ташхислаш алгоритмлари ва дастурий таъминотини яратиш» (2017-2019) мавзуларидаги лойиҳалари доирасида бажарилган.

Тадқиқотнинг мақсади ахборот ва қарор қабул қилишга кўмаклашиш тизими маълумотлар базалари учун интеллектуал таҳлилнинг самарали алгоритмларини ишлаб чиқиш ва сервисга йўналтирилган архитектурага асосланган дастурий восита моделларини лойиҳалаш.

Тадқиқотнинг вазифалари:

мавжуд ахборот ва қарор қабул қилишга кўмаклашиш тизимлари маълумотларини интеллектуал таҳлил қилиш усул, алгоритм ва дастурий воситаларни таҳлил қилиш;

маълумотларни интеллектуал таҳлил қилишнинг вақтли қаторларни башоратлаш, белгилар орасидаги боғлиқликларни аниқлаш ва ўхшаш объектларни саралаш каби масалалари учун самарали алгоритмлар ишлаб чиқиш;

сервисга йўналтирилган архитектура асосида маълумотларни интеллектуал таҳлил қилиш дастурий воситасининг функционал IDEF0 ва инфологик IDEF1x моделларини ишлаб чиқиш;

ахборот ва қарор қабул қилишга кўмаклашиш тизимлари ва маълумотларни интеллектуал таҳлил қилувчи дастурий восита ўртасида маълумотлар алмашиш тузилмасини ишлаб чиқиш;

интеллектуал таҳлил дастурий воситасини амалиётга тадбиқи, ахборот тизимларига интеграцияси, натижалар таҳлили ва техникавий жиҳозларга қўйиладиган талабларни ишлаб чиқиш.

Тадқиқотнинг объекти сифатида ахборот тизимларидаги маълумотлар базалари, қарор қабул қилишга кўмаклашишда маълумотларни интеллектуал таҳлил қилиш жараёнлари қаралган.

Тадқиқотнинг предмети ахборот ва қарор қабул қилишга кўмаклашиш тизимларини ишлаб чиқиш технологиялари, маълумотлар интеллектуал таҳлилининг усул ва алгоритмлари, сервисга йўналтирилган дастурий таъминотни лойиҳалаш воситаларидан иборат.

Тадқиқотнинг усуллари сифатида тадқиқотда тизими таҳлил, ахборот ва қарор қабул қилишга кўмаклашиш тизимларини ишлаб чиқиш, маълумотларнинг интеллектуал таҳлили, сервисга йўналтирилган дастурий таъминотларни ишлаб чиқиш, объектга йўналтирилган дастурлаш усулларида фойдаланилган.

Тадқиқотнинг илмий янгилиги қуйдагилардан иборат:

маълумотларни интеллектуал таҳлил қилишда ахборот ва қарор қабул қилишга кўмаклашиш тизимлари маълумотлар базалари учун вақтли қаторларни башоратлаш масаласини ечишнинг нейрон тўрлари асосида такомиллаштирилган алгоритм ишлаб чиқилган;

ахборот ва қарор қабул қилишга кўмаклашиш тизимлари маълумотлар базалари учун ўхшаш объектларни саралашнинг баҳоларни ҳисоблашга асосланган алгоритми ишлаб чиқилган;

ахборот ва қарор қабул қилишга кўмаклашиш тизимлари билан интеллектуал таҳлил қилувчи дастурий воситанинг маълумотлар алмашиш тузилмаси ишлаб чиқилган;

маълумотларнинг интеллектуал таҳлили масалаларини ечишга мўлжалланган самарали алгоритмларига асосланган дастурий воситанинг функционал IDEF0 ва инфологик IDEF1x моделлари ишлаб чиқилган.

Тадқиқотнинг амалий натижалари қуйдагилардан иборат:

ахборот тизимлари учун қарор қабул қилишга кўмаклашиш модулларини ишлаб чиқишда вақтли қаторларни башоратлаш, объект белгилари орасидаги боғлиқликларни аниқлаш ва ўхшаш объектларни саралаш масалаларини ечиш алгоритми ва дастури ишлаб чиқилган;

сервисга йўналтирилган архитектура асосида ахборот тизимлари учун интеллектуал таҳлил масалаларини ечишга мўлжалланган, ҳисоблашларни алоҳида серверда бажариб, муайян форматда натижани тақдим қилувчи веб сервис типидagi дастурий восита ишлаб чиқилган.

Тадқиқот натижаларининг ишончлилиги. Тадқиқот натижаларининг ишончлилиги қўйилган муаммонинг математик жиҳатдан коррект ифодаланиши, объект белгилари орасидаги боғлиқликларни аниқлашда, вақтли қаторларни башоратлашда ва ўхшаш объектларни саралашда маълумотларни интеллектуал таҳлил қилиш усулларининг тўғри қўлланилиши, шунингдек, назарий ва амалий тадқиқотлардан олинган натижалар ва уларнинг мувофиқлиги орқали изоҳланади.

Тадқиқот натижаларининг илмий ва амалий аҳамияти. Тадқиқот натижаларининг илмий аҳамияти тақлиф этилган ахборот тизимлари базасидаги маълумотлар асосида ўхшаш объектларни саралаш, белгилар орасидаги боғлиқликларни аниқлаш, вақтли қаторларни башоратлаш алгоритмлари ва дастурлараро маълумот алмашиш тузилмасини ишлаб чиқиш билан изоҳланади.

Тадқиқот натижаларининг амалий аҳамияти ахборот тизимларида маълумотларнинг кўпайиб кетиши натижасида юзага келадиган муаммоли вазиятларда, ишлаб чиқилган веб-сервис кўринишидаги дастурий восита билан ахборот тизими маълумотларини тезкор таҳлил қилишга мўлжалланган қарор қабул қилишга кўмаклашиш модулларини қўйиш иш самарадорлигини ошириши билан изоҳланади.

Тадқиқот натижаларининг жорий қилиниши. Ахборот тизимлари маълумотлари ёрдамида қарор қабул қилишга кўмаклашувчи модулларни

ишлаб чиқишга имкон берувчи интеллектуал таҳлил дастурий воситаси асосида:

беморлар оқимини башоратлаш, симптомлар орасидаги яширин боғлиқликларни аниқлаш ва ўхшаш ташхисларни саралаш каби қарор қабул қилишга қўмаклашиш модулларига эга ахборот тизими Республика ихтисослаштирилган кардиология маркази Хоразм филиалига жорий этилган. (Соғлиқни сақлаш вазирлигининг 2019 йил 13 майдаги 8н-4/157-сон ва Ахборот технологиялари ва коммуникацияларини ривожлантириш вазирлигининг 2018 йил 16 ноябрдаги 33-8/8597-сон маълумотномалари). Илмий тадқиқот натижаси аҳолига тиббий хизматларни кўрсатиш учун дастлабки ташхисларни тақдим этиш жараёнини оптималлаштириш имконини берган;

тиббий ахборот тизими маълумотлари асосида қарор қабул қилишга қўмаклашиш модуллари ишлаб чиқиш алгоритмлари Хоразм вилояти ихтисослашган урология марказига жорий этилган (Соғлиқни сақлаш вазирлигининг 2019 йил 13 майдаги 8н-4/157-сон ва Ахборот технологиялари ва коммуникацияларини ривожлантириш вазирлигининг 2018 йил 16 ноябрдаги 33-8/8597-сон маълумотномалари). Илмий тадқиқот натижаси тиббий хизматлар кўрсатиш жараёнини оптималлаштириш ҳамда беморларни касаллик ташхисларининг аниқлик даражасини 1,15 мартага ошириш имконини берган;

тиббий ахборот тизимларига беморлар келишини башоратлаш, ўхшаш ташхисларни саралаш, симптомлар орасидаги яширин боғлиқликларни аниқлаш каби модулларига эга интеллектуал таҳлил дастурий воситаси «Mustafo Software» МЧЖга жорий этилган (Ахборот технологиялари ва коммуникацияларини ривожлантириш вазирлигининг 2018 йил 16 ноябрдаги 33-8/8597-сон маълумотномаси). Мазкур дастурий воситадан фойдаланиш ахборот тизимларини ишлаб чиқишда иш жараёни самарадорлигини 1,25 мартага ошириш имконини берган.

Тадқиқот натижаларининг апробацияси. Диссертациянинг назарий ва амалий жиҳатлари 7 та халқаро ва 6 та республика илмий-амалий анжуманларида муҳокамадан ўтказилган.

Тадқиқот натижаларининг эълон қилинганлиги. Тадқиқот мавзуси бўйича 29 та илмий иш чоп этилган, шулардан, Ўзбекистон Республикаси Олий аттестация комиссияси тавсия этган илмий нашрларда 8 та мақола 7 таси республика ва 1 таси хорижий журналларда чоп этилган ҳамда 3 та ЭҶМ учун яратилган дастурий воситаларни қайд қилиш гувоҳномалари олинган.

Диссертациянинг тузилиши ва ҳажми. Диссертация кириш, тўртта боб, хулоса, фойдаланилган адабиётлар рўйхати ва иловалардан иборат. Диссертациянинг ҳажми 96 бетни ташкил этади.

ДИССЕРТАЦИЯНИНГ АСОСИЙ МАЗМУНИ

Кириш қисмида диссертация ишининг долзарблиги ва зарурияти асосланган, тадқиқот мақсади, вазибалари ҳамда объект, предмети

тавсифланган, Ўзбекистон Республикаси фан ва технологияси тараққиётининг устувор йўналишларига мослиги кўрсатилган, тадқиқотнинг илмий янгилиги, амалий натижалари баён қилинган, олинган натижаларнинг назарий ва амалий аҳамияти очиб берилган, тадқиқот натижаларини жорий қилиш рўйхати, нашр этилган ишлар ва диссертация тузилиши бўйича маълумотлар келтирилган.

Диссертациянинг «Қарор қабул қилишга қўмаклашиш тизимлари ва маълумотларнинг интеллектуал таҳлили» деб номланган биринчи бобида, ахборот ва қарор қабул қилишга қўмаклашиш тизимларининг тузилиши, таркибий қисмлари, маълумотлар интеллектуал таҳлилининг вақтли қаторларни башоратлаш, синфларга ажратиш, ассоциатив қоидаларни излаш масалаларининг таҳлили, маълумотларни интеллектуал таҳлил қилиш учун мавжуд дастурий воситаларнинг таҳлили ҳамда дастурий таъминот ишлаб чиқиш соҳасидаги замонавий дастурлаш тиллари, технологияларининг таҳлили ва масаланинг қўйилиши келтирилган.

Ҳозирда ахборот тизимларида маълумотлар ҳажмининг кўпайиб кетиши кўйидаги муаммоларни келтириб чиқармоқда:

катта ҳажмдаги маълумотлар орасидан керакли маълумотни саралаб олиш қийинлашади;

маълумотларни саралаб олишда, маълумотларга ишлов беришда кутиш ҳолатлари пайдо бўлади;

катта ҳажмли маълумотлар ҳолатида ахборот тизимларида фойдаланувчи томонидан амалга ошириладиган таҳлилларни ҳам ҳисоблаш тизимларига юклаш зарурати туғилади;

ахборот тизимларида катта ҳажмли маълумотларга статистик ва интеллектуал ишлов беришда катта ҳажмдаги ҳисоблашларни талаб қилади.

Шунинг учун ахборот тизимлари ишлаш самардорлигини янада яхшилаш мақсадида катта массивли маълумотларни қайта ишлаб, маълумотлар орасидаги боғлиқликларни таҳлил қилиб берадиган дастурий воситаларга, яъни қарор қабул қилишга қўмаклашиш модулларига эҳтиёж туғилади.

Мазкур бобда шунингдек мавжуд маълумотларни интеллектуал таҳлил қилувчи дастурий воситалари ва ахборот ва қарор қабул қилишга қўмаклашиш тизимларини ишлаб чиқишдаги жараёнлар ҳамда дастулаш технологиялари таҳлили ҳам келтириб ўтилган.

Худди шунингдек, ахборот ва қарор қабул қилишга қўмаклашиш тизимлари учун маълумотларни интеллектуал таҳлил қилишнинг усул ва алгоритмларини ўзида жамлаган, ҳисоблаш ишларини алоҳида серверда бажариб, ечимларни муайян форматда ахборот ва қарор қабул қилишга қўмаклашиш тизимларига тақдим қилувчи веб сервис дастурий воситасини ишлаб чиқиш асосий мақсад қилиб олинган. У кўйидаги масалаларни ечишни ўз ичига олади:

Фараз қилайлик, ахборот тизимида қайд қилиб бориш натижасида шакллантирилган маълумотлар базасидан олинган кўйидаги $\{S\}$ маълумотлар тўплами берилган бўлсин:

Объектлар	1	2	3	...	n	t
S_1	x_{11}	x_{12}	x_{13}	...	x_{1n}	td_1
S_2	x_{21}	x_{22}	x_{23}	...	x_{2n}	td_2
S_3	x_{31}	x_{32}	x_{33}	...	x_{3n}	td_3
...
S_m	x_{m1}	x_{m2}	x_{m3}	...	x_{mn}	td_m

Бу ерда S_i ($i = \overline{1, m}$) - объектлар ва x_{ij} ($i = \overline{1, m}, j = \overline{1, n}$) i -объектнинг j -чи белгиси, td_i ($i = \overline{1, m}$) - i -объект маълумотлар базасига киритилган сана.

Берилган маълумотлар тўплами асосида қуйидаги масалаларни ечиш белгилаб олинди:

- вақтли қаторларни башоратлаш;
- белгилар орасидаги боғлиқликларни аниқлаш;
- ўхшаш объектларни саралаш.

Диссертациянинг «Маълумотларни интеллектуал таҳлил қилишнинг самарали алгоритмлари» деб номланган иккинчи бобида вақтли қаторларда башоратлаш, белгилар орасидаги боғлиқликларни аниқлаш, ўхшаш объектларни саралаш масалаларини ечиш алгоритмлари келтириб ўтилган. Ишлаб чиқилган алгоритмлар диссертация ишида мақсад қилиб олинган веб-сервис дастурий воситасида фойдаланиш учун умумлаштирилган.

Мазкур бобнинг биринчи параграфи вақтли қаторларни башоратлаш масаласини ечиш учун алгоритм ишлаб чиқишга қаратилган. Масалани ечишда вақт қатори t – вектор компоненталари t_i , ($i=1,2,\dots,m$) $m \in Z$ кўринишида берилган бўлсин. Бу ерда t – вектор компоненталари td вектори компоненталари қийматларини саналар бўйича саралашдан ҳосил қилинади. Вектор компоненталарини «сурилувчи ойна» усули ёрдамида нейрон тўрлари учун ўқитилувчи танланмалар шаклига келтириб олинади. Агар ойнанинг ўлчами d га тенг бўлса, у ҳолда вақт қатори кетма-кетлигидаги вектор қуйидагича $\{T1\}$ объектлар тўплами кўринишга ўтади.

	киришлар	натижалар
	$t(1), t(2), t(3), \dots, t(d)$	$t(d+1)$
	$t(2), t(3), t(4), \dots, t(d+1)$	$t(d+2)$

	$t(m-d-1), t(m-d), t(m-d+1), \dots, t(m-1)$	$t(n)$

«Сурилувчи ойна» усули ёрдамида олинган маълумотлар тўпламини $\{T1\}$ ўқув танланма шаклига келтириб олинади ва у қуйидагича бўлади.

Объектлар	1	2	3	...	d	$d+1$
$T1_1$	$t1_{11}$	$t1_{12}$	$t1_{13}$...	$t1_{1d}$	ty_1
$T1_2$	$t1_{21}$	$t1_{22}$	$t1_{23}$...	$t1_{2d}$	ty_2
...
$T1_m$	$t1_{m1}$	$t1_{m2}$	$t1_{m3}$...	$t1_{m1d}$	ty_{m1}

Бу ерда $m1=n-d-1$, «сурилувчи ойна» усули ёрдамида олинган

маълумотлар тўпламини $\{TI\}$ ўқув танланмадаги объектлар сони $ty_i, (i = \overline{1, m1}) - \{TI\}$ ўқув танланмадаги натижалар вектори, $t1_{ij}, (i = \overline{1, m1}, j = \overline{1, d}) - \{TI\}$ ўқув танланмадаги объектлар параметрларининг кийматлари.

Ҳосил қилинган $\{TI\}$ ўқув танланма маълумотлари асосида уч қатламли тўғри тақсимланган нейрон тўри қуриб олинади.

Уч қатламли тўғри тақсимланган нейрон тўри $IN=d, HN=d, ва ON=1$ да самарали бўлади, бу ерда IN - кириш қатламидаги нейронлар сони, HN - яширин қатламдаги нейронлар сони, ON - чиқувчи қатламдаги нейронлар сони. d соҳа маълумотларининг даврийлигини инобатга олган ҳолда танлаш мақсадга мувофиқ.

Шунга асосан уч қатламли тўғри тақсимланган нейрон тўрига асосланган вақтли қаторларни башоратлаш масаласини ечиш алгоритми қуйидагича бўлади:

1-қадам: Берилган вақтли қатор маълумотлари $t_i, i = 1, m$ векторга юклаб олинади.

2-қадам: «Сурилувчи ойна»нинг ўлчами d ва тест натижалари сони l киритилади.

3-қадам: «Сурилувчи ойна» усули ёрдамида t_i вектордан ўқитилувчи танлангани характерловчи $t1_{ij}$ массив ҳосил қилинади.

4-қадам: Уч қатламли тўғри тақсимланган нейрон тўри ҳосил қилинади. Бунда кирувчи қатламдаги нейронлар сони $nl=d$, яширин қатламдаги нейронлар сони $kl=d$, чиқиш қатламдаги нейронлар сони $m1=1$, $W_{ij} (i, j = 1, nl+k1+m1)$ вазнлар массиви тасодиқий сонлар билан инициализация қилинади.

5-қадам: Ўқитилувчи танланмадан $t1_{ij}$ киришлар вектори уч қатламли тўғри тақсимланган нейрон тўри берилади ва

$a_k = \sum_{k1=1}^d t1_{ik1} * w_{k1,k}$ формулага асосан a_k ҳисобланади $g(a) = \frac{1}{1-e^{-a}}$ формулага асосан ва ty' лар ҳисобланади.

6-қадам: $E(w) = \frac{1}{2} \sum_{i=1}^{m1} (ty_i - ty'_i)$ формула билан чиқиш қатлами хатоликлари E_k ҳисобланади.

7-қадам: $\delta_j = g'(a_j) \sum_k \delta_k w_{jk}$ (1) формула билан барча яширин қатламлар хатоликлари δ_j ҳисобланади.

8-қадам: $\frac{\partial E}{\partial w_{ji}} = \delta_j z_i$, формула бўйича жорий градиент ҳисобланади.

9-қадам: $\Delta w_{ji} = -\eta \frac{\partial E}{\partial w_{ji}}$ формула билан вазнлар массиви ўзгартирилади,

яъни $w(n+1) = w(n) + \Delta w$ Бу ерда η - ўқитиш тезлиги $0 < \eta < 1$

10-қадам: (1) формулага асосан хатолик киймати ҳисобланади. Агар хатолик етарлича минимал бўлса, ҳисоблаш тўхтатилади ва 11-қадамга ўтилади, акс ҳолда 5-қадамга ўтилади.

11- қадам: Олинган W_{ij} ёрдамида берилган t_i векторнинг $m+1, m+2, \dots, m+l+l$ элементлари башорат қилинади.

Мазкур алгоритм асосида ишлаб чиқилган дастур биз диссертацияда мақсад қилиб олган дастурий воситанинг модули бўлиб хизмат қилади.

Мазкур бобнинг иккинчи параграфида белгилар орасидаги боғлиқликларни аниқлаш алгоритми ишлаб чиқилган:

Бобнинг 3-параграфида ўхшаш объектларни саралаш алгоритми тимсолларни таниб олишнинг баҳоларни ҳисоблаш алгоритмига асосланган ҳолда ишлаб чиқилган:

1. *Таянч тўплamlар тизими.* Бу босқичда берилган S маълумотлар тўпламидан x_{ij} матрица кўринишидаги таянч тўплaml ҳосил қилинади.

2. *Яқинлик функцияси.* Бу босқичда таянч тўплaml объектларининг белгилари асосида ҳар бир белги учун яқинлилик параметри қуйидагича ҳисобланади:

$$\rho_j(x_{ij}, x_{kj}) = \begin{cases} 1, & \text{агар } |x_{ij} - x_{kj}| \leq \varepsilon_j \\ 0, & \text{агар } |x_{ij} - x_{kj}| > \varepsilon_j \end{cases}, j = \overline{1, n}; i, k = \overline{1, m} \quad (2)$$

$$\varepsilon_j = \frac{1}{m-1} \sum_{i=1}^{m-1} |x_{ij} - x_{i+1j}|, j = \overline{1, n} \quad (3)$$

3. *Фиксирланган таянч тўплamlнинг қаторлари бўйича баҳоларни ҳисоблаш.* Бу босқичда янги келган $S_{m+1} = (x_{m+11}, x_{m+12}, \dots, x_{m+1n})$ объектга нисбатан ҳар бир объектнинг овозлари ҳисобланади.

$$\Gamma_i = \Gamma_i(S_i, S_{m+1}) = \sum_{j=1}^n \rho_j(x_{ij}, x_{m+1j}), i = \overline{1, m} \quad (4)$$

$$\Gamma_{j_1} \geq \Gamma_{j_2} \geq \dots \geq \Gamma_{j_n}$$

Мазкур ҳисоблаш босқичларига асосланган ўхшаш объектларни саралаш алгоритми қуйидагича бўлади:

1-қадам: S маълумотлар тўплamlи объектларни x_{ij} массивга юклаб олинади.

2-қадам: яқинлик функцияси (2) формула ва бўсағавий қийматлар (3) формула орқали ҳисобланади.

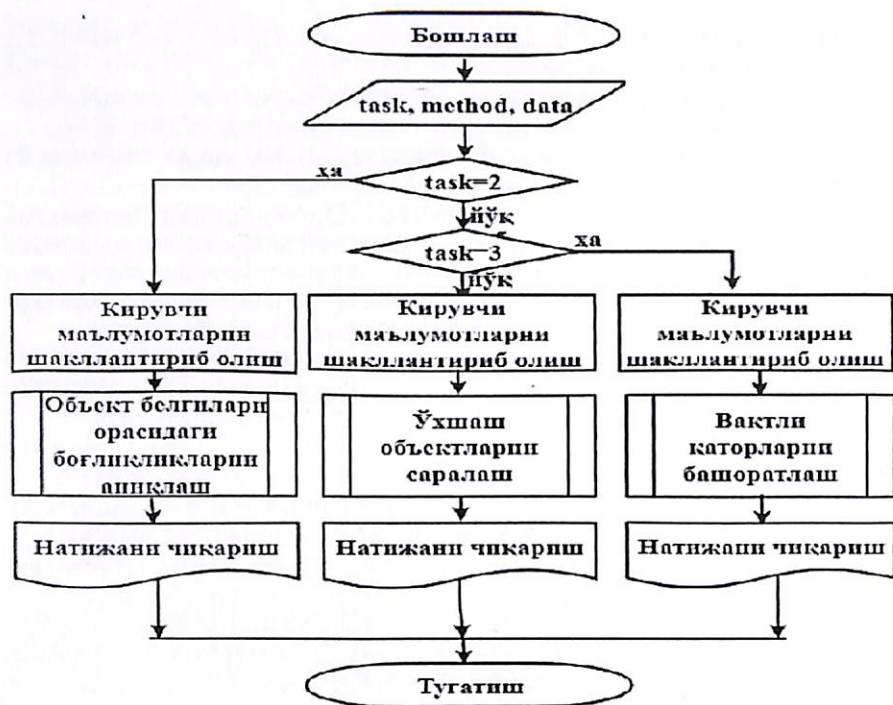
3-қадам: Янги S_{m+1} объект x_{m+1j} кўринишида x_{ij} массивга юкланади ва кидирилаётган ўхшаш объектлар сони v ни киритилади.

4-қадам: ҳар бир объектнинг янги объектга нисбатан олган овозлари (4) формула орқали ҳисобланади ҳамда o_i векторда юкланади.

5-қадам: оддий саралаш усули ёрдамида x_{ij} массив o_i векторга асосан камайиш тартибида сараланади.

6-қадам: x_{ij} массивнинг дастлабки v та сатри чоп қилинади.

Мазкур бобнинг тўртинчи параграфида олдинги учта параграфда кўрилган учта масала учун ишлаб чиқилган алгоритмлар умумлаштирилади ва қуйидагича алгоритм ишлаб чиқилади. Алгоритмнинг блок-схема кўриниши 1-расмда келтирилган.



1-расм. Умумий алгоритмининг блок-схема кўриниши.

Диссертациянинг учинчи боби «Сервисга йўналтирилган архитектура асосида маълумотларни интеллектуал таҳлил қилиш дастурий воситасини ишлаб чиқиш» деб номланган бўлиб, унда ҳозирги кун дастурлаштиришидаги энг оммабоп технологиялардан бири ҳисобланган сервисга йўналтирилган дастурий таъминот архитектураси, унинг зарурати, таркибий қисмлари ва унинг асосида ишлаб чиқилган дастурий воситалар, иккинчи бобда кўриб чиқилган вактли каторларни саралаш, белгилар орасидаги боғлиқликларни аниқлаш, ўхшаш объектларни саралаш масалалари учун танланган алгоритмларни ўзида жамлаштирган «DM-API» (Data Mining Application Programming Interface) дастурий воситасининг архитектураси, унинг IDEF0 функционал ва инфологик IDEF1.x модели, унга бўлган талаблар, ишлаб чиқиш босқичлари ва уни тестлаш каби масалалар кўриб чиқилган.

Бобнинг биринчи параграфида «DM-API» дастурий воситасинининг IDEF0 модели келтирилган (2-расм):



2-расм. DM-API дастурий воситасининг IDEF0 функционал модели.

Кирувчи маълумотларга қуйидагилар киради:

Вақтли каторлар асосида башоратлаш масаласини ечиш учун вектор кўринишидаги чекли сондаги катор маълумотлар;

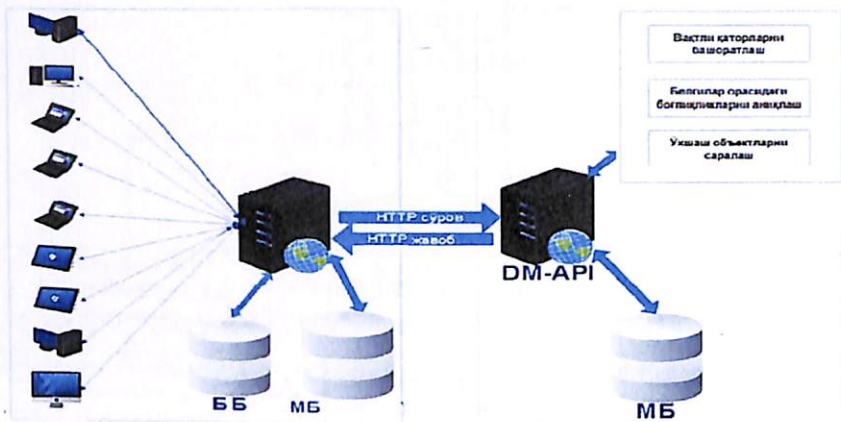
Белгилар орасидаги боғлиқликларни аниқлаш учун сонлардан иборат матрица кўринишидаги маълумотлар;

Ўхшаш объектларни саралаш учун сонлардан иборат матрица кўринишидаги маълумотлар;

Усуллар сифатида эса уч қатламли тўғри тақсимланган нейрон тўри модели, белгилар орасидаги боғлиқликларни аниқлаш усули, ўхшаш объектларни саралаш усуллари кириб, улар моделга таъсир қилади.

Моделга таъсир қилувчи мезонлар сифатида хатоликларни камаййтириш, яқинлилик ва объектга берилган баҳоаларни келтириш, шунингдек, моделдан чиқувчи натижа сифатида интеллектуал таҳлил натижаларини келтириш мумкин.

Дастурий воситанинг ахборот тизимлари билан маълумот алмашиш тамойили 3-расмда келтирилган.



3–расм. «DM-API» дастурий воситасининг маълумот алмашиш тамойили.

Ахборот тизимларини «DM-API» дастурий воситаси билан маълумот алмашиш мақсадида қуйидагича тузилма ишлаб чиқилди. Унинг умумий кўриниши қуйидагича:

$Structure = \langle token, task, method, data \rangle$

token- веб сервисдан фойдаланувчи регистрацидан ўтиш даврида унга бериладиган уникал калит. Шу орқали веб сервис ким маълумот жўнатаётганлигини ва унга жавоб қайтариш зарур ёки зарурмаслигини билиб олади.

task- бу веб сервис қандай масалани ечиш лозимлигини аниқлашчи параметр.

method- бу *task* ўзгарувчисига боғлиқ массив кўринишида бўлиб асосан масалага қараб турлича элементлардан ташкил топади. Бу тузилма орқали масала ечилаётган алгоритм параметрлари узатилади.

data- асосий маълумотларни ўзида сақловчи икки ўлчовли массив кўринишида бўлиб ундаги маълумотлар ечиладиган масаланинг турига қараб турлича кўринишда бўлади.

Бобнинг учинчи параграфида «DM-API» дастурий воситасининг инфологик IDEF 1x модели ва дастурий воситадан фойдаланиб олинган натижалар асосида билимлар базасинини шакллантириш қаралган.

Бобнинг тўртинчи параграфида ишлаб чиқилган «DM-API» дастурий воситадан фойдаланиш учун аппарат ва дастурий таъминотга қўйилган минимал талаблар ишлаб чиқилган.

Диссертациянинг ««DM-API» дастурий воситасидан олинган натижалар таҳлили ва уни ахборот тизимларига тадбиқи» деб номланган тўртинчи бобида дастурий воситаси ёрдамида «E-Cardio» ахборот тизимида қарор қабул қилишга кўмаклашиш модулларини қўлланилиши кўриб

чиқилади.

«E-Cardio» ахборот тизими маълумотлар базасидаги маълумотлардан интеллектуал таҳлил жараёни учун маълумотлар витринаси қуриб олиб маълумотлар омборига сақлаб олинади. Бунда бемор маълумотлари n та атрибутдан иборат S объектлар массиви сифатида юборилади. Мазкур объектлар таҳлили асосида «E-Cardio» ахборот тизимига «DM-API» дастурий воситаси орқали қуйидаги модуллар қўшилди.

Беморлар оқимини башоратлаш:

Мазкур модулга «E-Cardio» ахборот тизими маълумотлар базасидаги уч йиллик беморларнинг кундалик таширфларнинг сони асосида шакллантирилган 1093 та объект билан кейинги ой учун беморлар оқимини башоратлаш масаласи ечилди ва натижалар олинди. Натижалар 4-расмда келтирилган.



4-расм. Кейинги ой учун башоратлаш натижалари

4-расмдаги кейинги ой учун башоратлаш натижалар шунини кўрсатадики, нейрон тўрига асосланган модель бу масала учун энг яхши натижа кўрсатган. Бу модулдан фойдаланиш ихтисослашган кардиология маркази бош шифокорига кейинги ойга келадиган беморларнинг сонини янада аниқроқ башорат қилишга ва шунга кўра режалаштириш ишларини олиб боришида муҳим аҳамият касб этади.

Симптомлар орасидаги яширин боғлиқликларни топиш:

Мазкур алгоритм ёрдамида кардиология марказидаги лаборатория текширувларидан бўлган ВЭМ, ЭХО, ЭКГ ва қон таҳлили натижаларидан 21 та белгидан ва 119 та объектдан иборат маълумотлар тўплами танлаб олинди (1-жадвал).

Натижадан кўриш мумкинки, белгилар орасидаги боғлиқликлар алгоритми ёрдамида маълумотлар тўплами таҳлил қилинган ва энг ишончли 10 та қоида келтириб чиқарилган. Шунингдек натижалар таҳлилида, 119 та объектдан 86 тасининг 4- ва 5- белгилари бир хил эки 4-белги ва 6- белгиларнинг қиймати 5-белгига 84

«E-Cardio» ахборот тизими маълумотлари асосидаги белгилар орасидаги боғлиқликлар натижалари.

№	Қаралаётган белгилар	Такрорланишлар сони	Ишончлилик кўрсаткичи
1	$x_4=1.4$ ва $x_5=4.4$	86	1
2	$x_4=1.4$, $x_6=35.8$ ва $x_5=4.4$	84	1
3	$x_6=35.8$ ва $x_5=4.4$	87/86	0.99
4	$X_4=1.4$ ва $x_6=35.8$	86/84	0.98
5	$x_5=4.4$, $x_6=35.8$ ва $x_4=84$	86/84	0.98
6	$x_4=1.4$ $x_5=4.4$ ва $x_6=35.8$	86/84	0.98
7	$x_5=4.4$ ва $x_4=1.4$	89/86	0.97
8	$x_5=4.4$ ва $x_6=35.8$	89/86	0.97
9	$x_6=35.8$ ва $Col_4=1.4$	87/84	0.97

Мазкур натижаларга кўра, қуйидагича хулосалар олиш мумкин:

1) Беморлар лаборатория текширувлари натижаларидаги 4-симптом билан 5-симптом жуда яқин, чунки мазкур симптомлар 119 та ҳолатдан 86 тасида ўхшаш натижаларни такрорлаган.

2) Тўртинчи, олтинчи ва бешинчи симптомлар 84 мартаба бир хил кийматларни қабул қилган. Демак, бу симптомлар орасидаги боғлиқликлар юқори.

Хулоса қилиб айтганда, ахборот тизимларидаги маълумотлари устида белгилар орасидаги боғлиқликларни аниқлаш соҳа мутахассисларига муҳим қарорларни чиқаришда ёрдам беради.

Ўхшаш таъхисларни саралаш:

Тажриба маълумотлари сифатида Республика ихисослашган кардиология маркази Хоразм вилояти бўлими «E-Cardio» ахборот тизими маълумотлар базасидаги 90 нафар беморнинг 6 та лаборатория текширувларидан 21 та белги бўйича берилган маълумотларидан фойдаланилди.

2-жадвалда келтирилган дастур натижаларидан кўриш мумкинки, жами овозларни ҳисоблаш учун барча комбинациялар 352716 та, натижада овоз олишлар бўйича 1-ўринни 48-объект эгаллаган бўлиб, у тўплаган овозлар сони 187756 та, шу сабабли мазкур объект синов объектимизга энг яқин объект ҳисобланади. Шундай тарзда бизга зарур бўлган L та синов объектимизга энг яқин объектларни топишимиз мумкин. Бу ўз навбатида бир нечта беморларнинг маълумотлари орасидан шифокорга керак бўлган беморлар натижаларини тез топишга ёрдам беради.

Тажриба натижалари қуйидагича:

Дастлаб яқинлилик вектори ҳисоблаб олинади (2-жадвал).

Ўхшаш объектларни саралаш алгоритми натижалари

Белгилар	1	2	3	4	5	6	7	8	9	10	11
Яқинлилик вектори	7,20	0,43	133,27	0,05	0,02	0,43	2,75	4,71	16,77	55,08	21,35
Тестланувчи объект	150,00	5,02	5,40	1,40	4,40	35,80	16,00	2,20	58,00	296,00	25,80
1-энг яқин объект	147,00	4,98	7,00	1,44	4,40	35,80	7,00	1,80	68,00	325,00	29,60
2-энг яқин объект	111,00	4,04	6,80	1,40	4,40	35,80	10,00	2,30	58,00	286,00	21,80
3-энг яқин объект	113,00	4,06	5,40	1,40	4,40	35,80	7,00	1,30	66,00	325,00	29,60
4-энг яқин объект	127,00	4,71	8,00	1,40	4,40	35,80	9,00	1,70	72,00	154,00	29,60
Белгилар	12	13	14	15	16	17	18	19	20	21	
Яқинлилик вектори	0,83	0,97	55,45	41,37	12,61	0,41	0,17	0,18	1,83	1,05	
Тестланувчи объект	5,40	4,10	140,00	76,00	46,00	1,00	0,90	3,50	3,50	1,70	
1-энг яқин объект	5,60	3,70	153,00	58,00	61,00	1,00	1,00	3,60	4,00	2,20	
2-энг яқин объект	5,60	4,40	154,00	89,00	42,00	1,40	1,00	3,60	3,40	2,50	
3-энг яқин объект	5,80	4,50	164,00	91,00	44,00	1,10	1,00	3,50	4,30	1,90	
4-энг яқин объект	5,20	3,80	130,00	64,00	50,00	1,40	1,00	3,60	4,20	1,60	

Мазкур модулда шифокорлар томонидан янги келган беморнинг лаборатор текширувлардан ўтган натижалари ва аввалги маълумотлар базасида мавжуд беморларнинг натижалари модул томонидан «DM-API» дастурий воситасига юборилади. Натижада дастурий восита эса тизимга маълумотлар базасида мавжуд беморларнинг натижалари орасидан янги бемор натижаларига тўртта энг яқин натижаларни саралаб жўнатади.

ХУЛОСА

«Қарор қабул қилишга қўмаклашиш тизимлари учун интеллектуал таҳлил дастурий воситасини ишлаб чиқиш» мавзусидаги диссертация бўйича олиб борилган тадқиқотлар натижасида қуйидаги хулосалар тақдим этилди:

1. Маълумотларни интеллектуал таҳлиллаш дастурий воситалари, уларнинг мақсад ва вазибаларига ҳамда дастурий таъминотларни ишлаб чиқиш стандартларига мослиги, маълумотларга дастлабки ишлов бериш, ахборот тизимларига интеграцияси каби хоссалари кесимида таҳлил қилинди. Ўрганилган дастурий воситалар хоссалари кесимида ишлаб чиқилган

таҳлилий жадвал, дастурий воситаларни солиштирма таҳлил қилиш ва уларни мавжуд ахборот тизимларига интеграциясини аниқлашга имкон берди.

2. Вақтли қаторларни башоратлаш масаласини ечиш учун уч қатламли тўғри тақсимланган нейрон тўри моделига асосланган такомиллашган алгоритми ишлаб чиқилди. Бунинг натижасида ахборот тизимлари маълумотлар базаларидаги вақтга боғлиқ равишда сақланган маълумотлардан фойдаланган ҳолда кейинги бўлиши мумкин бўлган маълумотларни башоратлаш имконини берди.

3. Ахборот тизимлари объектларининг белгилари орасидаги боғлиқликларни аниқлашнинг алгоритми ишлаб чиқилди. Ишлаб чиқилган алгоритмдан фойдаланган ҳолда ахборот тизимлари маълумотлар базаларида йиғилган маълумотлардан дастлабки ишлов бериш асосида аниқланган белгилар орасидаги боғлиқликлар таҳлил қилиниб, соҳа ходимларига муҳим хулосалар бериш имконини берди.

4. Ўхшаш объектларни саралашнинг баҳоларни ҳисоблашга асосланган алгоритми ишлаб чиқилди. Бунинг натижасида ахборот тизимлари маълумотлар базасидаги маълумотлар кўпайиб таҳлил қилишни мураккаблаштирган вазиятларда ўхшаш объектларни саралаб олиб объектларни таҳлил қилиш имконини берди.

5. «DM-API» дастурий воситасига кирувчи маълумотлар, унга таъсир килувчи усуллар, шу усулларнинг таъсир мезонлари, дастурий воситадан натижа сифатида чиқувчи натижавий маълумотлар аниқлаштирилди ва шулар асосида дастурий воситанинг IDEF0 функционал модели, IDEF0 функционал моделнинг декомпозицияси ва инфологик IDEF1x модели ишлаб чиқилди. Бунинг натижасида ишлаб чиқиши режалаштирилган «DM-API» дастурий воситасининг модуллари, унинг таркибий қисмлари, маълумотлар базасининг тузилиши ва унда амалга ошириладиган ҳисоблаш амалларини аниқлаштириб олиш имконияти пайдо бўлди.

6. «DM-API» дастурий воситасининг модуллари, тарихий қисмлари ва ундан фойдаланиш тадқиқ қилинди ва ахборот тизимлари билан «DM-API» дастурий воситаси ўртасида маълумот алмашиш тузилмаси ишлаб чиқилди. Бунинг натижасида, ахборот тизимлари учун қарор қабул қилишга кўмаклашиш модулларини ишлаб чиқишда «DM-API» дастурий воситаси билан маълумот алмашиш имконияти пайдо бўлди.

7. Ишлаб чиқилган «DM-API» дастурий воситасидан фойдаланиш учун аппарат ва дастурий жиҳатдан зарурий талаблар ишлаб чиқилди. Тизим икки қисмдан иборат эканлигини эътиборга олган ҳолда, ҳам мижоз, ҳам сервер тарафдан дастурий воситадан фойдаланиш бўйича аппарат ва дастурий таъминотларга бўлган талаблар алоҳида ишлаб чиқилди. Бу талабларнинг ишлаб чиқиши дастурий воситадан фойдаланиш жараёнида келиб чиқиши мумкин бўлган хатолик ва муаммоларни бартараф этиш имконини берди.

**ТАШКЕНТСКИЙ УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ
ТЕХНОЛОГИЙ**

ХУЖАЕВ ОТАБЕК КАДАМБАЕВИЧ

**РАЗРАБОТКА ПРОГРАММНОГО СРЕДСТВА
ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДЛЯ СИСТЕМ ПОДДЕРЖКИ
ПРИНЯТИЯ РЕШЕНИЙ**

05.01.04 – Математическое и программное обеспечение вычислительных машин,
комплексов и компьютерных сетей

**АВТОРЕФЕРАТ ДИССЕРТАЦИИ
ДОКТОРА ФИЛОСОФИИ (PhD) ПО ТЕХНИЧЕСКИМ НАУКАМ**

ВВЕДЕНИЕ (аннотация диссертации доктора философии (PhD))

Актуальность и востребованность темы диссертации. В мире тенденция развития индустрии по производству программных продуктов требует особого внимания информационным систем и систем поддержки принятия решений. В 2016 году только оборот рынка от программного обеспечения по обработке электронных карт пациентов в медицинской сфере составил 20,55 миллиарда долларов. «На сегодняшний день программное средство медицинской информационной системы «Open EMR» реализовано в более чем 10 государствах и в более чем 1500 больницах»¹. Особенно в США, Германии, Франции, Великобритании, Японии, Австралии, Южной Кореи, Китае, Индии, Российской Федерации и Узбекистане уделяется большое внимание разработке программного обеспечения информационных систем и систем принятия поддержки решений, применяемых для развития различных сфер общества, а также разработке специальных программных средств для обслуживания этих систем.

В мире проводится большое количество научных исследований для повышения эффективности работы в соответствующих сферах деятельности путем применения информационных систем и систем принятия поддержки решений, то есть путем интеллектуального анализа больших массивов данных в базе данных информационных систем, выявления зависимостей между их признаками, сортировки аналогичных объектов, прогнозирования временных рядов. В том числе, разработка программных средств поддержки принятия решения, которые обрабатывают данные в системах управления базами данных, методами и алгоритмами интеллектуального анализа данных является одной из актуальных задач.

В республике особое внимание уделяется улучшению качества медицинского обслуживания, в частности, повышению качества услуг в области кардиологии на основе информационно-коммуникационных технологий и средств компьютерной диагностики. В стратегии действий по пяти приоритетным направлениям развития Республики Узбекистан на 2017-2021 годы определены задачи, в том числе «... прежде всего первичного звена, скорой и экстренной медицинской помощи, направленное на повышение доступности и качества медицинского и социально-медицинского обслуживания населения, формирование здорового образа жизни населения, укрепление материально-технической базы медицинских учреждений, ... внедрение информационно-коммуникационных технологий в социальную сферу, системы управления»². Реализация этих задач, в частности, в сфере кардиологии применение программных модулей интеллектуального анализа к информационным системам и тем самым повышение эффективности деятельности врачей является одной из важнейших задач.

Диссертационное исследование в определенной мере послужит

¹ http://open-emr.org/wiki/index.php/OpenEMR_Professional_Support

² Указ Президента Республики Узбекистан от 7 февраля 2017 г. №УП-4947 «О Стратегии действий по дальнейшему развитию Республики Узбекистан»

выполнению задач, предусмотренных указом Президента Республики Узбекистан от 7 февраля 2017 года № УП-4947 «О стратегии действий по дальнейшему развитию Республики Узбекистан», Постановлением Президента от 2 августа 2018 года № ПП-3894 «О мерах по внедрению инновационной модели управления здравоохранением в Республике Узбекистан», Постановлением Президента от 21 марта 2012 года № ПП-1730 «О мерах по дальнейшему внедрению и развитию современных информационно-коммуникационных технологий» и Постановлением Кабинета Министров от 1 февраля 2012 года № 24 «О мерах по созданию условий для дальнейшего развития компьютеризации и информационно-коммуникационных технологий на местах», а также и других нормативно-правовых актов, касающихся данной сферы деятельности.

Соответствие исследования приоритетным направлениям развития науки и технологий республики. Данное исследование выполнено в рамках приоритетных направлений развития науки и технологии Республики IV. «Информатизация и развитие информационно-коммуникационных технологий».

Степень изученности проблемы. Многие ученые проводили научные исследования в области систем поддержки принятия решений, информационные системы, интеллектуальный анализ данных, распознавания образов, проектирования и разработки сервисно-ориентированного программного обеспечения, разработка и использования методов обработки данных и математическое моделирование.

К примеру, в работах С.С. Aggarval, E.S. Bernera изучены информационные системы и системы принятия поддержки решений, модели их построения, в работах С.М. Bishop, R.O. Duda, P.E. Hart, А.А. Барсегян, В.Н. Вапник, В.И. Васильев, Ю.И. Журавлев, Н.Г. Загоруйко, Г.С. Лбов, Л.А. Растргина рассмотрены задачи распознавания образов, прогнозирования временных рядов и классификация объектов при интеллектуальном анализе данных.

Так же ученые и исследователи нашей Республики М.М. Камилов, Т.Ф. Бекмуратов, Ш.Х. Фазылов, Р.Х. Хамдамов, М.А. Рахматуллаев, Н.А. Игнатьев, Д.Т. Мухамедиева, А.Х. Нишанов, Р.Н. Усманов, А.Р. Ахатов, А.И. Назаров внесли свой вклад в научные исследования в этом направлении.

До сегодняшнего дня недостаточно исследован вопрос разработки программных средств на основе сервисно-ориентированной архитектуры и предоставления решений для различных платформ в целях интеллектуального анализа данных информационных систем.

Связь диссертационного исследования с планами научно-исследовательских работ высшего образовательного учреждения, где выполнена диссертация. Диссертационное исследование выполнено в рамках проектов плана научно-исследовательских работ Ташкентского университета информационных технологий имени Мухаммада ал-Хоразмий: ФА-А17-Ф056 «Алгоритмическое обеспечение адаптивной распознающей

системы интеллектуального анализа данных в условиях неопределенности информации об объектах» (2009-2012); БА-5-017 «Создание алгоритмов и программного обеспечения диагностики рака молочной железы и шейки матки на основе методов интеллектуального анализа данных» (2017-2019).

Целью исследования является разработка эффективных алгоритмов интеллектуального анализа базы данных информационной системы и системы поддержки принятия решений, проектирование моделей программных средств на основе сервисно-ориентированной архитектуры.

Задачи исследования:

анализ методов, алгоритмов и программных средств интеллектуального анализа данных информационной системы и системы поддержки принятия решений;

разработка эффективных алгоритмов для задач прогнозирования временных рядов, определения зависимостей между признаками и сортировки схожих объектов в интеллектуальном анализе данных;

разработка инфологической IDEF1x и функциональной IDEF0 моделей программного средства интеллектуального анализа на основе сервисно-ориентированной архитектуры;

разработка структуры обмена информацией между программным средством интеллектуального анализа данных и информационной системой и системой поддержки принятия решений;

внедрение программных средств интеллектуального анализа, интеграция с информационными системами, анализ результатов и разработка требований к техническим устройствам.

Объектом исследования являются базы данных информационных систем, процессы интеллектуального анализа данных при поддержке принятия решений.

Предмет исследования – технология разработки информационных систем и систем поддержки принятия решений, методы и алгоритмы интеллектуального анализа данных, средства проектирования сервисно-ориентированных программных обеспечений.

Методы исследования включают системный анализ, разработку информационных систем и систем поддержки принятия решений, интеллектуальный анализ данных, разработку сервисно-ориентированного программного обеспечения, методы объектно-ориентированного программирования.

Научная новизна исследования заключается в следующем:

при интеллектуальном анализе данных на основе нейронных сетей разработаны усовершенствованные алгоритмы решения задачи прогнозирования временных рядов для базы данных информационной системы и системы поддержки принятия решений;

разработан алгоритм сортировки аналогичных объектов для базы данных информационной системы и систем поддержки принятия решений, основанный на вычислении оценок;

разработана структура обмена информацией между программным средством интеллектуального анализа и информационной системой;

разработаны функциональная IDEF0 и инфологическая IDEF1x модели программного средства, которые включают в себе эффективные алгоритмы решения задач интеллектуального анализа данных.

Практические результаты исследования заключаются в следующем:

при создании вспомогательных модулей принятия решений для информационной системы разработаны алгоритм и программа решения задач прогнозирования временных рядов, определения зависимостей между признаками объекта и сортировки схожих объектов;

на основе сервисно-ориентированной архитектуры разработано веб-платформенное программное средство, которое решает задачу интеллектуального анализа данных информационной системы и выполняющее вычисления в отдельном веб-сервере, а также представляющее результаты в определенном формате.

Достоверность результатов исследования обосновывается корректным математическим выражением поставленной проблемы, правильным применением методов интеллектуального анализа данных при определении зависимостей между признаками объектов, при прогнозировании временных рядов, при сортировке схожих объектов, а также соответствием теоретических и практических полученных результатов исследования.

Научная и практическая значимость результатов исследования. Научная значимость исследования разъясняется алгоритмами сортировки схожих объектов, определения зависимостей между признаками и прогнозирования временных рядов на основе базы данных предложенной информационной системы, а также разработанной структурой обмена данными между программами.

Практическая значимость результатов исследования заключается в том, что в проблемных ситуациях, возникающих с увеличением данных в информационных системах, программное средство в виде веб-сервиса помогает повысить производительность работы информационной системы с помощью вспомогательных модулей поддержки принятия решений ориентированных для оперативного анализа данных.

Внедрение результатов исследования. На основе программного средства интеллектуального анализа, которое дает возможность разработки вспомогательных модулей принятия решения на основе данных информационных систем:

информационная система с внедренными модулями поддержки принятия решений для определения скрытой взаимосвязи между симптомами и сортировки схожих диагнозов внедрена в Хорезмском филиале Республиканского специализированного кардиологического центра (справки Министерство здравоохранения от 13 мая 2019 г., № 8н-4/157 и Министерства по развитию информационных технологий и коммуникаций Республики Узбекистан от 16 ноября 2018 г. №33-8/8597). Результаты исследования дают

возможность оптимизировать процесс предоставления первичных диагнозов при медицинском обслуживании населения;

алгоритмы разработки вспомогательных модулей по поддержке принятия решений на основе данных медицинской информационной системы внедрено в Специализированном урологическом центре Хорезмской области (справка Министерства здравоохранения от 13 мая 2019 г., № 8и-4/157 и Министерства по развитию информационных технологий и коммуникаций Республики Узбекистан от 16 ноября 2018 г. №33-8/8597). Результаты исследования дают возможность оптимизировать процесс оказания медицинских услуг и повысить точность диагностирования заболеваний пациентов в 1,15 раза;

программное средство интеллектуального анализа с модулями прогнозирования поступления пациентов в медицинскую информационную систему, сортировка схожих диагнозов, определение скрытых зависимостей между симптомами внедрено в ООО «Mustafo Software» (справка Министерства по развитию информационных технологий и коммуникаций Республики Узбекистан от 16 ноября 2018 г. №33-8/8597). Применения упомянутых программных средств при разработке информационных систем программа дает возможность увеличить производительность рабочего процесса в 1,25 раза.

Апробация результатов исследования. Теоретические и практические аспекты диссертации обсуждены на 7 международных и на 6 республиканских научно-практических конференциях.

Публикация результатов исследования. По теме исследования опубликовано всего 29 научных работ, в том числе 8 статей в научных изданиях, рекомендованных Высшей аттестационной комиссией к публикации основных научных результатов докторских диссертаций, из них 1 статья в зарубежном журнале и 7 в республиканских журналах, а также получены свидетельства на 3 программные разработки для ЭВМ.

Структура и объём диссертации. Диссертация состоит из введения, четырех глав, заключения, списка использованной литературы и приложений. Объем диссертации составляет 96 страницы.

ОСНОВНОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ

Во введении диссертационной работы обоснованы актуальность и значимость темы, определены цели и задачи исследования, а также объект и предметы исследования, соответствие исследования приоритетным направлениям развития науки и технологий Республики Узбекистан, приведены научная новизна и практические результаты исследования, раскрыты теоретические и практические значения полученных результатов, приведены список внедрения результатов исследования, сведения об опубликованных работах, а также структура диссертации.

В первой главе диссертации «Системы поддержки принятия решений и интеллектуальный анализ данных» изложена структура, компоненты информационных систем и систем поддержки принятия решений, анализ задач

прогнозирования временных рядов, классификация, поиск ассоциативных правил при интеллектуальном анализе данных, анализ существующих программных средств для интеллектуального анализа данных, а также анализ современных языков и технологий в индустрии разработки программного обеспечения и приведена постановка задачи.

В настоящее время увеличение объема данных в информационных системах вызывает следующие проблемы:

усложнение сортировки и извлечение необходимой информации среди больших объемов данных;

появляется состояния ожидания при сортировке и обработки данных;

в случае больших объемов данных возникает необходимость проводить пользовательский анализ данных на базе вычислительной системы;

требуется большая вычислительная мощность информационной системы при статистической и интеллектуальной обработке больших объемов данных.

Поэтому в целях повышения эффективности работы информационных систем возникает потребность в программных средствах, то есть в модулях поддержки принятия решений, которые анализируют взаимосвязь между данными путем обработки больших массивов данных.

В упомянутой главе также приведен анализ существующих программных средств интеллектуального анализа данных, процессов разработки информационных систем и систем поддержки принятия решения, а также технологий программирования.

Точно также, выбрана как основная цель разработка программного обеспечения веб-сервиса, которая основана на методах и алгоритмах интеллектуального анализа данных информационных систем и систем поддержки принятия решений, а также предоставляет решения для информационных систем и систем поддержки принятия решений в определенном формате путем выполнения вычислительных работ на отдельном сервере. Она включает в себя решение следующих задач:

Предположим, дано набор данных $\{S\}$, которые сформированы в ходе фиксирования в базе данных информационной системы, в следующем виде:

Объекты	1	2	3	...	n	t
S_1	x_{11}	x_{12}	x_{13}	...	x_{1n}	td_1
S_2	x_{21}	x_{22}	x_{23}	...	x_{2n}	td_2
S_3	x_{31}	x_{32}	x_{33}	...	x_{3n}	td_3
...
S_m	x_{m1}	x_{m2}	x_{m3}	...	x_{mn}	td_m

Здесь S_i ($i = \overline{1, m}$) – объекты и x_{ij} ($i = \overline{1, m}, j = \overline{1, n}$) j -ый признак i -го объекта, td_i ($i = \overline{1, m}$) – i -дата введения объекта в базу данных.

На основе заданных наборов данных намечено решить следующие задачи:

- прогнозирование временных рядов;
- определение взаимосвязи между признаками;

– сортировка схожих объектов.

Во второй главе диссертации «Эффективные алгоритмы интеллектуального анализа данных» приведены алгоритмы прогнозирования временных рядов, определения взаимосвязей между признаками, решения задач сортировки схожих объектов. Разработанные алгоритмы обобщены для использования в программном обеспечении веб-сервиса, являющего целью диссертационной работы.

Первый параграф этой главы предусматривает разработку алгоритма прогнозирования временных рядов. Пусть при решении задачи компоненты t – вектора временного ряда представлены в виде $t_i, (i=1,2,\dots,m) m \in Z$. Здесь компоненты t – вектора образуются при сортировке значений компонентов td вектора по дате. Векторные компоненты при помощи метода «скользящего окна» приводятся к форме обучающих выборок для нейронных сетей. Если размер окна равен d , то вектор в последовательности временных рядов переходит в вид следующего набора объектов $\{Tl\}$.

входящие	результаты
$t(1), t(2), t(3), \dots, t(d)$	$t(d+1)$
$t(2), t(3), t(4), \dots, t(d+1)$	$t(d+2)$
.	.
.	.
.	.
$t(m-d-1), t(m-d), t(m-d+1), \dots, t(m-1)$	$t(n)$

Полученные наборы данных методом «скользящего окна» приводим к форме обучающих выборок $\{Tl\}$ и имеем следующий вид:

Объекты	1	2	3	...	d	$d+1$
Tl_1	t_{111}	t_{112}	t_{113}	...	t_{11d}	ty_1
Tl_2	t_{121}	t_{122}	t_{123}	...	t_{12d}	ty_2
...
Tl_{ml}	t_{1m11}	t_{1m12}	t_{1m13}	...	t_{1m1d}	ty_{m1}

Здесь $ml=n-d-1$ – число объектов набора данных, полученных методом «скользящего окна» в $\{Tl\}$ обучающей выборке; $ty_i, (i = \overline{1, ml})$ – вектор результатов в обучающей выборке $\{Tl\}$; $t_{1ij}, (i = \overline{1, ml}, j = \overline{1, d})$ – значения параметров объектов в обучающей выборке $\{Tl\}$.

На основе полученных данных обучающей выборки $\{Tl\}$ строится трехслойная нейронная сеть прямого распространения.

Трехслойная нейронная сеть прямого распространения, при $IN=d, HN=d$, и $ON=1$ будет эффективной, здесь, IN – количество нейронов во входном слое, HN – количество нейронов в скрытом слое, а ON – количество нейронов в выходящем слое. Обычно целесообразно выбирать d с учетом периодичности данных.

На основании этого алгоритм решения задачи прогнозирования временных рядов на основе трехслойной нейронной сети прямого

распространения состоит из следующих шагов:

Шаг 1. Данные временного ряда $t_i, i = 1, m$ загружаются в вектор.

Шаг 2. Вводятся d – размер «Скользящего окна» и l – количество тестовых результатов.

Шаг 3. С помощью метода «Скользящего окна» от вектора t_i образуется t_{1ij} массив, характеризующее обучающую выборку.

Шаг 4. Образуется трехслойная нейронная сеть прямого распространения. При этом количество нейронов во входном слое равно $n_l = d$, в скрытом слое – $k_l = d$, в выходном слое – $m_l = l$, массив весов $W_{ij} (i, j = 1, n_l + k_l + m_l)$ инициализируется случайными числами

Шаг 5. Вектор входов t_{1ij} от обучающей выборки передается трехслойной нейронной сети прямого распространения и на основе формулы $a_k = \sum_{k=1}^d t_{1ik1} * w_{k1,k}$ вычисляется a_k , и по формуле $g(a) = \frac{1}{1-e^{-a}}$ вычисляется ty' .

Шаг 6. По формуле $E(w) = \frac{1}{2} \sum_{i=1}^{m_l} (ty_i - ty'_i)^2$ вычисляются погрешности E_k выходного слоя.

Шаг 7. Формулой $\delta_j = g'(a_j) \sum_k \delta_k w_{jk}$ (1) вычисляются погрешности δ_j скрытых слоев.

Шаг 8. Формулой $\frac{\partial E}{\partial w_{\mu j}} = \delta_j z_j$ вычисляется текущий градиент.

Шаг 9. Формулой $\Delta w_{\mu j} = -\eta \frac{\partial E}{\partial w_{\mu j}}$ преобразуется матрица весов, то есть $w(n+1) = w(n) + \Delta w$. Здеc η – скорость обучения, $0 < \eta < 1$

Шаг 10. На основе формулы (1) определяются значения погрешности. Если погрешность достаточно минимальная, то вычисление останавливается и осуществляется переход к шагу 11, в противном случае осуществляется переход к шагу 5.

Шаг 11. При помощи полученной W_{ij} прогнозируется $m_l + 1, m_l + 2, \dots, m_l + l$ элементы t_i заданного вектора.

Программа, разработанная на основе приведенного алгоритма, служит модулем программного средства, который является целью диссертации.

Во втором параграфе данной главы описан алгоритм определения взаимосвязи между признаками.

В 3-параграфе данной главы алгоритм сортировки похожих объектов разработан, основываясь на алгоритме вычисления оценок распознавания образов:

1. Система опорных множеств. На этом этапе формируется опорное множество в виде x_{ij} матрицы от множеств данных S .

2. Функция близости. На данном этапе на основе признаков объектов опорного множества параметр близости для каждого признака рассчитывается следующим образом:

$$\rho_j(x_{ij}, x_{kj}) = \begin{cases} 1, & \text{если } |x_{ij} - x_{kj}| \leq \varepsilon_j \\ 0, & \text{если } |x_{ij} - x_{kj}| > \varepsilon_j \end{cases}, j = \overline{1, n}; i, k = \overline{1, m} \quad (2)$$

$$\varepsilon_j = \frac{1}{m-1} \sum_{i=1}^{m-1} |x_{ij} - x_{i+1j}|, j = \overline{1, n} \quad (3)$$

3. Расчет оценок по рядам зафиксированной опорной множеством. На данном этапе вычисляются голоса каждого объекта относительно вновь поступающего $S_{m+1} = (x_{m+1,1}, x_{m+1,2}, \dots, x_{m+1,n})$ объекта.

$$\Gamma_i = \Gamma_i(S_i, S_{m+1}) = \sum_{j=1}^n \rho_j(x_{ij}, x_{m+1j}), i = \overline{1, m} \quad (4)$$

$$\Gamma_{j_1} \geq \Gamma_{j_2} \geq \dots \geq \Gamma_{j_n}$$

Алгоритм сортировки похожих объектов на основе этих вычислительных этапов состоит из следующих шагов:

Шаг 1. Объекты множеств данных S загружаются в x_{ij} массив;

Шаг 2. Функция близости рассчитывается по формуле (2) и пороговым значениям по формуле (3).

Шаг 3. Загрузка S_{m+1} нового объекта в виде $x_{m+1,j}$ в массив x_{ij} и введение v количество разыскиваемых похожих объектов.

Шаг 4. Полученные голоса каждого объекта относительно нового объекта рассчитываются по формуле (4) и загружаются в вектор o_i

Шаг 5. На основе вектора o_i методом простой сортировки сортируется x_{ij} массив по убыванию.

Шаг 6. Печатается первая v строка x_{ij} массива.

В четвертом параграфе этой главы обобщаются алгоритмы, разработанные для трех задач в предыдущих трех параграфах, и разрабатывается алгоритм следующим образом. Вид алгоритма представлен на рис. 1.

В третьей главе диссертации «Разработка программного обеспечения интеллектуального анализа данных на основе сервисно-ориентированной архитектуры», рассматривается архитектура программного обеспечения, ориентированная на сервис, которая в настоящее время является одной из самых популярных технологий в программировании, ее необходимость, компоненты и разработанные на ее основе программные средства, сортировка временных рядов, рассмотренная во второй главе, определение взаимосвязей между признаками, архитектура программного обеспечения «DM-API» (Data Mining Application Programming Interface), в которой собраны выбранные алгоритмы для решения проблемы сортировки похожих объектов, включает в себя функциональная IDEF0 и инфологическая IDEF1.x модели, требования к ней, этапы разработки и тестирование.

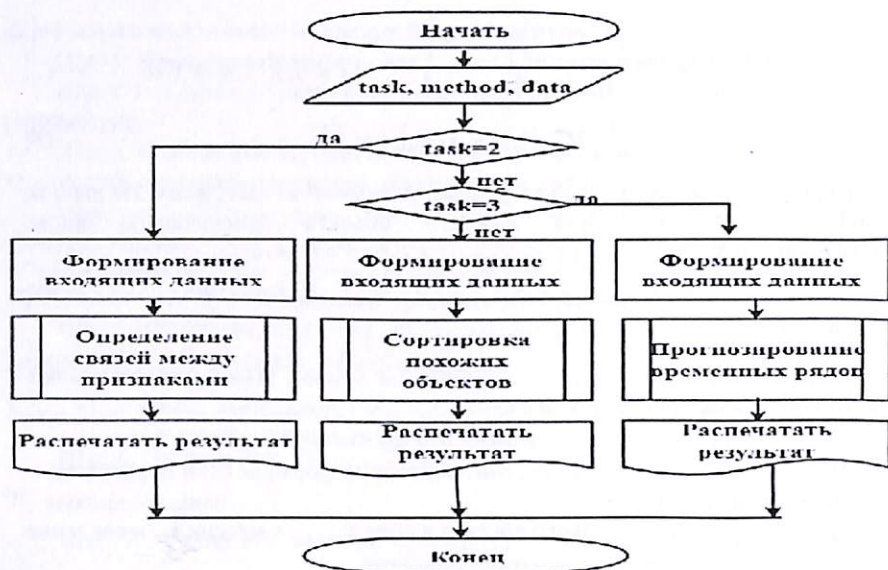


Рис. 1. Блок-схема обобщенного алгоритма

В первом параграфе главы представлена IDEF0 модель программного средства «DM-API» (Рис.2):

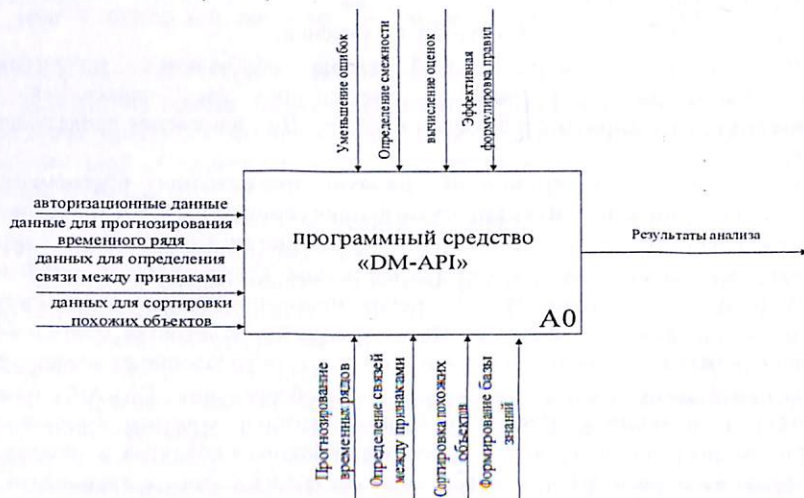


Рис. 2. Функциональная IDEF0 модель программного обеспечения «DM-API»

Входными данными являются:

Ряд данных ограниченного количества в векторном виде для решения задачи прогнозирования на основе временных рядов;

Данные в виде матрицы, состоящей из чисел, для определения взаимосвязи между признаками;

Данные в виде матрицы, состоящей из чисел для сортировки похожих объектов.

А в качестве методов входят модель трехслойной нейронной сети прямого распространения, определение связи между признаками, метод поиска похожих объектов, и они влияют на модель.

В качестве критериев, влияющих на модель, можно привести уменьшение ошибок, близость и оценка объектов, а также в качестве выходящего результата от модели результаты интеллектуального анализа.

Принцип обмена информацией с информационными системами программного обеспечения приведен на рис. 3.

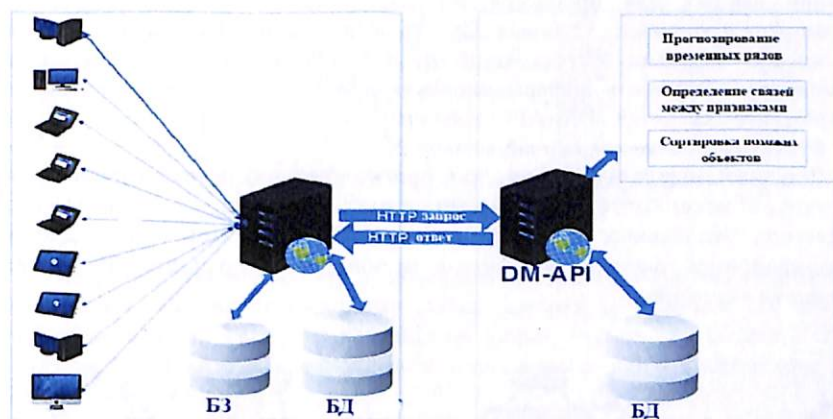


Рис. 3. Принцип обмена информацией программного обеспечения «DM-API»

Для обмена информацией программного обеспечения «DM-API» с информационными системами разработана следующая структура. Ее общий вид выглядит следующим образом:

Structure = <token, task, method, data>.

token- уникальный ключ, который предоставляется пользователю при регистрации в веб-сервисе. Таким образом, веб-сервис узнаёт, кто отправляет информацию, и должен ли он отвечать на нее или нет.

task- параметр, который означает, какую задачу должен решать веб-сервис.

method- имеет форму массива, зависящей от переменной *task*, состоит из

различных элементов в зависимости от решаемой задачи. Через эту структуру передаются параметры алгоритма решаемой задачи.

data- двумерный массив, содержащий в себе основные данные, которые зависят от типа решаемой задачи.

В третьем параграфе главы рассмотрены инфологическая модель IDEF1x программного средства «DM-API» и формирования базы знаний на основе полученных результатов с использованием программного средства.

В четвертом параграфе главы разработаны минимальные требования к аппаратному и программному обеспечению для использования разработанного программного средства «DM-API».

В четвертой главе диссертации «Анализ полученных результатов с помощью программного обеспечения «DM-API» и его внедрение в информационные системы» рассматривается применение модулей поддержки принятия решений в информационной системе «E-Cardio» с помощью программного средства.

На основе базы данных информационной системы «E-Cardio» создается витрина данных для процессов интеллектуального анализа данных и сохраняется в хранилище данных. При этом данные пациента отправляются как массив объектов S , состоящий из n атрибутов. На основе анализа упомянутых объектов в информационную систему «E-Cardio» посредством программного средства «DM-API» добавлены следующие модули.

Прогнозирование потока пациентов

В данном модуле решается вопрос прогнозирования потока пациентов на следующий месяц с 1093 объектами, которые сформированы на основе трехлетних ежедневных посещаемых количества пациентов в базе данных информационной системы «E-Cardio» и получены результаты. Результат показан на 4-рисунке.



Рис. 4. Результаты прогнозирования на следующий месяц

Прогнозы на следующий месяц на рисунке 4 показывают, что модель, основанная на нейронной сети, показала лучший результат для этой задачи.

Главному врачу специализированного кардиологического центра использование этого модуля может иметь важное значения для более точного прогнозирования количества прибывающих пациентов на следующий месяц и планирования работы.

Нахождение скрытых взаимосвязей между симптомами

С помощью данного алгоритма из результатов лабораторных исследований ВЭМ, ЭХО, ЭКГ и анализа крови в кардиологическом центре выбран набор данных, состоящий из 21 признака 119 объекта (таблица 1).

Таблица-1.

Результаты взаимосвязи между признаками полученных данных от информационной системы E-Cardio

№	Рассматриваемые признаки	Количество повторений	Показатель надежности
1	$x_4=1.4$ и $x_5=4.4$	86	1
2	$x_4=1.4$, $x_6=35.8$ и $x_5=4.4$	84	1
3	$x_6=35.8$ и $x_5=4.4$	87/86	0.99
4	$x_4=1.4$ и $x_6=35.8$	86/84	0.98
5	$x_5=4.4$, $x_6=35.8$ и $x_4=84$	86/84	0.98
6	$x_4=1.4$ $x_5=4.4$ и $x_6=35.8$	86/84	0.98
7	$x_5=4.4$ и $x_4=1.4$	89/86	0.97
8	$x_5=4.4$ и $x_6=35.8$	89/86	0.97
9	$x_6=35.8$ и $Col4=1.4$	87/84	0.97

Из результата видно, что с помощью алгоритма взаимосвязей между признаками был проанализирован набор данных и получено 10 самых надежных правил. Также из результатов видно, что из 119 объектов в 86 объектах признаки 4- и 5- идентичны или признаки 4- и 6- взаимосвязаны с 5-ым признаком 84 раз.

По этим результатам можно сделать следующие выводы:

- 1) в результате лабораторных исследований пациентов с 4 симптом очень близок к симптому 5, поскольку эти симптомы повторили аналогичные результаты в 86 случае из 119;
- 2) четвертый, шестой и пятый симптомы получили одинаковые значения 84 раза, а это значит, что связь между этими симптомами высокая;

Таким образом, выявление взаимосвязей между признаками данных в информационных системах поможет специалистам сферы при принятии важных решений.

Сортировка схожих диагнозов

В качестве экспериментальных данных использованы данные 6 лабораторных исследований для 90 больных по 21 признаку в базе данных информационной системы «E-Cardio» Хорезмского областного отделения Республиканского специализированного кардиологического центра.

По результатам программы, приведенной на таб.2 видно, что для подсчета голосов в общей сложности все комбинации равны 352716, в результате 1-ое место по голосованию занял 48-ой объект с количеством набранных голосов - 187756, поэтому данный объект является самым близким к испытываемому объекту. Таким образом, мы можем найти необходимые нам объекты, которые ближе всего к нашему L-ому испытываемому объекту. Это, в свою очередь, помогает быстро найти необходимые для врача результаты пациентов среди нескольких пациентов.

Результаты эксперимента следующие:

Таблица-2.

Результаты алгоритма сортировки схожих объектов

Признаки	1	2	3	4	5	6	7	8	9	10	11
Вектор интервалов	7,20	0,43	133,27	0,05	0,02	0,43	2,75	4,71	16,77	55,08	21,35
Тестируемый объект	150,00	5,02	5,40	1,40	4,40	35,80	16,00	2,20	58,00	296,00	25,80
ближайший объект 1	147,00	4,98	7,00	1,44	4,40	35,80	7,00	1,80	68,00	325,00	29,60
ближайший объект 2	111,00	4,04	6,80	1,40	4,40	35,80	10,00	2,30	58,00	286,00	21,80
ближайший объект 3	113,00	4,06	5,40	1,40	4,40	35,80	7,00	1,30	66,00	325,00	29,60
ближайший объект 4	127,00	4,71	8,00	1,40	4,40	35,80	9,00	1,70	72,00	154,00	29,60
Признаки	12	13	14	15	16	17	18	19	20	21	
Вектор интервалов	0,83	0,97	55,45	41,37	12,61	0,41	0,17	0,18	1,83	1,05	
Тестируемый объект	5,40	4,10	140,00	76,00	46,00	1,00	0,90	3,50	3,50	1,70	
ближайший объект 1	5,60	3,70	153,00	58,00	61,00	1,00	1,00	3,60	4,00	2,20	
ближайший объект 2	5,60	4,40	154,00	89,00	42,00	1,40	1,00	3,60	3,40	2,50	
ближайший объект 3	5,80	4,50	164,00	91,00	44,00	1,10	1,00	3,50	4,30	1,90	
ближайший объект 4	5,20	3,80	130,00	64,00	50,00	1,40	1,00	3,60	4,20	1,60	

В этом модуле со стороны врачей результаты лабораторных исследований вновь прибывшего пациента и результаты предыдущих пациентов в базе данных модулем отправляются в программное средство «DM-API». В результате программное средство сортирует данные пациентов из базы данных и как результат отправляет первые 4 самые близкие данные пациентов к данным вновь прибывшего пациента.

ВЫВОДЫ

В результате исследования по диссертации на тему «Разработка программного средства интеллектуального анализа для систем поддержки

принятия решений» представлены следующие выводы:

1. Программные средства интеллектуального анализа данных анализировались в разрезе таких характеристик, как соответствие их целям и задачам, а также стандартам разработки программного обеспечения, первичная обработка данных, интеграция с информационными системами. Аналитическая таблица, разработанная в разрезе свойств изученных программных средств, позволила провести сравнительный анализ программных средств и определить их интеграцию в существующие информационные системы.

2. Для решения проблемы прогнозирования временных рядов разработан усовершенствованный алгоритм, основанный на модели трехслойной нейронной сети прямого распространения. В результате этого, используя хранящуюся зависимую от времени информацию в базах данных информационных систем, появилась возможность прогнозировать вновь поступающую информацию.

3. Разработан алгоритм определения взаимосвязей между признаками объектов информационной системы. С помощью разработанного алгоритма была проанализирована взаимосвязь между данными в базах данных информационной систем на основе предварительной обработки, и выявлена возможность получить важные выводы сотрудникам отрасли.

4. Разработан алгоритм сортировки схожих объектов, основанный на вычислениях оценок. В результате этого появилась возможность сортировать и анализировать подобные объекты в тех случаях, когда анализ данных в базе данных информационной системы усложнено из-за большой размерности данных.

5. Определены данные, входящие в программное средство «DM-API» и влияющие на него методы, критерии воздействия этих методов, результирующие данные, выходящие из программного средства, на основе которых была разработана IDEF0 функциональная модель программного средства, декомпозиция функциональной модели IDEF0 и инфологическая модель IDEF1x. В результате этого появилась возможность выявления модулей, компонентов запланированной для разработки программного средства «DM-API», структуры базы данных и осуществляемых в ней вычислительных операций.

6. Исследованы модули, структура и использование программного обеспечения «DM-API», а также разработана структура обмена данными между информационными системами и программным средством «DM-API». В результате появилась возможность обмена данными программным средством «DM-API» с информационными системами при разработке модулей систем поддержки принятия решений.

7. Разработаны необходимые требования к аппаратным и программным средствам для использования разработанного программного средства «DM-API». Учитывая то, что система состоит из двух частей, требования к аппаратной и программной средствам разработаны отдельно для клиентской

и серверной части для использования программного средства. Разработка этих требований позволила устранить ошибки и проблемы, которые могут возникнуть в процессе использования программного средства.

KHUJAEV OTABEK KADAMBAEVICH

**DEVELOPMENT OF A SOFTWARE TOOL FOR INTELLECTUAL
ANALYSIS FOR DECISION SUPPORT SYSTEMS**

05.01.04 – Mathematical and software support of computers,
complexes and computer networks

**DISSERTATION ABSTRACT OF THE DOCTOR OF PHILOSOPHY (PhD) ON
TECHNICAL SCIENCES**

The theme of doctor of philosophy (PhD) on technical sciences was registered at the Supreme attestation commission at the Cabinet of Ministers of the Republic of Uzbekistan under number B2017.4.PhD/T53.

The dissertation has been prepared at the Tashkent University of Information Technologies.

The abstract of the dissertation is posted in three languages (Uzbek, Russian, English (resume)) on the website www.tuit.uz and on the website of «ZiyoNet» Information and educational portal www.ziyounet.uz.

Scientific adviser: Nishanov Axram Hasanovich
Doctor of Technical Sciences

Official opponents: Muminov Bahodir Boltaevich
Doctor of Technical Sciences

Khaitov Fayzulla Norbutaevich
Candidate of Technical Sciences, Docent

Leading organization: Tashkent State Technical University

The defense will take place « 14 » June 2019 at 14⁰⁰ the meeting of Scientific council No. DSc.27.06.2017.T.07.01 at Tashkent University of Information Technologies (Address: 100202, Tashkent city, Amir Temur street, 108. Tel.: (+99871) 238-64-43, fax: (+99871) 238-65-52, email: tuit@tuit.uz).

The dissertation can be reviewed at the Information Resource Centre of the Tashkent University of Information Technologies (is registered under No. 580). (Address: 100202, Tashkent city, Amir Temur street, 108. Tel.: (+99871) 238-64-43, fax: (+99871) 238-65-52).

Abstract of dissertation sent out on « 01 » June 2019 y.
(mailing report No. 9 on « 02 » May 2019 y.).



R.Kh.Khamdamov
Chairman of the scientific council
awarding scientific degrees,
Doctor of Technical Sciences, Professor

F.M.Nuraliev
Scientific secretary of scientific council
awarding scientific degrees,
Doctor of Technical Sciences, Docent

H.N.Zaynidinov
Chairman of the academic seminar under the
scientific council awarding scientific degrees,
Doctor of Technical Sciences, Professor

INTRODUCTION (abstract of PhD dissertation)

The aim of the research work is to develop the efficient algorithms for the information system and decision support system, projecting software models based on a service-oriented architecture.

The object of the research work is the databases of information systems, database processes to support decision making.

The scientific novelty of the research work is consist on follows:

algorithms of modification based on neural networks have been developed for solving the problem of time series forecasting for database on information system and a decision support system;

developed algorithm based on the calculation of estimates the similar objects for the database on the information system and decision support system;

structure has been developed for the exchange of information between the software for intellectual analysis and the information system, and the decision support system;

functional IDEF0 and infological IDEF1x software models have been developed, which include efficient algorithms for solving intellectual analysis tasks.

Implementation of the research results. On the basis of a software tool for intellectual analysis, which enables the development of auxiliary decision-making modules based on data from information systems of specialized cardiological centers:

information system "E-Cardio" was implemented into Khorezm Branch of the Republican Specialized Cardiological Center (certificates of Ministry of Health of the Republic of Uzbekistan dated May 13, 2019 No. 8H-4/157 and Ministry for Development of Information Technologies and Communications of the Republic of Uzbekistan dated November 16, 2018 No. 33-8 / 8597). The result of research makes it possible to optimize the process of presenting forecasts of health care;

on the basis of data of medical information system of development algorithms for auxiliary decision-making modules introduced in the Specialized Urology Center of the Khorezm region (certificates of Ministry of Health of the Republic of Uzbekistan dated May 13, 2019 No. 8H-4/157 and Ministry for Development of Information Technologies and Communications of the Republic of Uzbekistan dated November 16, 2018 No. 33-8 / 8597). The result of research makes it possible to improve the accuracy of diseases diagnoses and the process of health care up to 1.15 times;

software "DM-API", which includes modules for diagnosing patients for admission to medical information system, sorting similar diagnoses based on laboratory results of disease classification, the definition of hidden dependencies between symptoms was implemented in "Mustafo Software" LLC (certificate of Ministry for Development of Information Technologies and Communications of the Republic of Uzbekistan dated November 16, 2018 No. 33-8 / 8597). While developing information systems, the program provides an opportunity to increase the productivity of the workflow up to 1,25 times.

Structure and volume of dissertation. Dissertation consists of introduction, four chapters, conclusion, list of references and applications. The dissertation volume is 96 pages.

ЭЪЛОН ҚИЛИНГАН ИШЛАР РЎЙХАТИ
СПИСОК ОПУБЛИКОВАННЫХ РАБОТ
LIST OF PUBLISHED WORKS

1. Khujaev O.K., Allamov O.T., Finding effective architecture of neural networks for time series prediction // Кимёвий технология. Назорат ва бошқарув. №4. –Тошкент, 2013.-Р. 82-89. (05.00.00; №12).

2. Рўзибоев О.Б., Хўжаев О.К., У.А.Хасанов Таянч векторлар усули асосида классификациялаш хатоликлари тахлили // «Informatika va energetika muammolari» O'zbekiston jurnali. №3-4.-Toshkent, 2013.–Б. 82-89.(05.00.00; №5).

3. Рўзибоев О.Б., Хўжаев О.К. Исследование и программная реализация метода ближайших соседей // Кимёвий технология. Назорат ва бошқарув. №2. –Ташкент, 2014. –Б. 84-89. (05.00.00; №12).

4. Нишанов А.Х., Акбаралиев Б.Б., Рўзибоев О.Б., Хўжаев О.К. Сравнительный анализ алгоритмов на основе нечеткого метода k-средних с применением различных метрик // Кимёвий технология. Назорат ва бошқарув. №6. –Ташкент, 2014. –С. 78-82. (05.00.00; №12).

5. Акбаралиев Б.Б., Рўзибоев О.Б., Хўжаев О.К., Саидрасулов Ш.Н. Moodle тизими базаси элементлари кесимида data mining усулларидан фойдаланган ҳолда ўқитувчи фаолиятини баҳолаш // ТАТУ хабарлари. №2(34). -Ташкент, 2015. – Б. 11-17. (05.00.00; №31).

6. Хўжаев О.К., Саидрасулов Ш.Н., Алламов О.Т. Маълумотларни интеллектуал тахлили дастурий воситаларининг қўлланиш соҳалари // Muhammad al-Xorazmiy avlodlari. №1(3). -Ташкент, 2018.-Б. 98-101. (05.00.00; №10).

7. Khujaev O.K. Selection of architecture and training algorithms of neural networks for classification solutions // European science review № 3-4. Austria – Viena -2018. –Р. 120-124. (05.00.00; №3).

8. Хўжаев О.К., Эгамбердиев Н.А. Тиббиёт соҳаси ахборот тизимларида ўқшаш ташхисларни топиш учун баҳоларни ҳисоблаш алгоритми қўллаш усули // Muhammad al-Xorazmiy avlodlari. №4(6).-Ташкент, -2018. -Б. 15-18. (05.00.00; №10).

9. Хўжаев О.К., Абдуллаев Ф.О., Рахманова М.Р. “Методы определение веса документов в электронных ресурсах” // “Science and world” International scientific journal. № 3 (19) Vol I. -Volgograd, -2015. –С. 10-14.

10. Хўжаев О.К. Абдуллаев Ф.О., Артиков М.Э., Бобожанов Б. Применение нейронных сетей в поисковых машинах типа электронного правительства // “Science and world” International scientific journal. № 3 (19) Vol I. –Volgograd, -2015. –С 14-18.

11. Хўжаев О.К., Артиков М.Э., Абдуллаев Ф.О., Худайберганов Т.А. Компарирование методов классификации в задачах медицинского диагностирования // “Science and world” International scientific journal. № 5 (21), Vol I. –Volgograd, -2015. –С. 10-13.

12. Хўжаев О.К., Курбанова О.У. Intellectual tahlil masalalari uchun weka

dasturiy ta'minoti // «Молодой учёный» международный научный журнал. № 9.5 (113.5). –Казань,-2016. –В. 39-40.

13. Хўжаев О.К., Садуллаев Н.Д., Klasterizatsiya masalalarini yechishda optimal algoritmi tanlash // «Молодой учёный» международный научный журнал. № 9.5 (113.5). –Казань,-2016. –В. 32-35.

14. Хўжаев О.К., Садуллаев Н.Д., Weka API yordamida klasterizatsiya masalasini yechish // «Ученый XXI века» научный журнал. № 6-3. Йошкар-Ола-2016. –В.11-14.

15. Хўжаев О.К., Алламов О.Т. Кўп қатламли нейрон тўрида ўқитиш алгоритмларини танлаш // «Ахборот технологиялари ва телекоммуникация муаммолари» Республика илмий-техник конференцияси маърузалар тўплами. –Тошкент,-2012.-Б. 132-134.

16. Kuryazov D.A., Khujaev O.K., Allamov O.T. Evolution of integrated software models // Proceedings Seventh World Conference on Intelligent Systems for Industrial Automation. Volume I. -Tashkent-2012. 25-27 November. -P. 165-169.

17. Khujaev O.K., Paredes R.P. Search optimal architecture of neural networks for time series prediction // Proceedings Seventh World Conference on Intelligent Systems for Industrial Automation. Volume I. Tashkent-2012. 25-27 November. -P. 173-178.

18. Хўжаев О.К., Алламов О.Т., Рўзиев О.Б. Маълумотлар интеллектуал тахлилидаги эркин ва очик кодли дастурий таъминотлар хақида // «Ахборот технологиялари ва телекоммуникация муаммолари», Республика илмий-техник анжумани маърузалар тўплами, 1-қисм. Ташкент-2013. –Б. 89-91.

19. Хўжаев О.К. Виталий П.С., Ядгаров Ш. Сравнение моделей нейронной сети для прогнозирования временных рядов // «Технические науки от теории к практике» сборник статей по материалы XXIV международный научно-практический конференции. №6(31).-Новосибирск, -2014. –С. 22-30.

20. Ruzibayev O.B., Asrayev M.A., Khujaev O.K. Comparing different data mining techniques for medical data classification // International scientific conference “Perspectives for development of information technologies ITPA-2014”. Tashkent-2014. -P. 222-226.

21. Хўжаев О.К., Рузиев О.Б., Асраев М.А., Исследование метода нечетких с-средних используя среду MATLAB // «Амалий математика ва информацион технологияларнинг долзарб муаммолари-ал-Хоразмий 2014» конференция материаллари. Тўплам №2, -Самарқанд, -2014. -С.146-150.

22. Khujaev O.K., Ruzbayev O.B., Allamov O.T., Vrachlar faoliyatini yaxshilash uchun CDSS (clinical decision support system) lar va ularning metodologik asoslari haqida // «Амалий математика ва информацион технологияларнинг долзарб муаммолари-АЛ-ХОРАЗМИЙ 2014» халқаро илмий конференцияси. Том №2. -Самарқанд, -2014. -В.117-119.

23. Boboev L.B., Khujayev O.K., Ruzibayev O.B. Comparative analyzing features selection methods for data mining tasks // Proceedings eighth world conference on intelligent systems for industrial automation - WCIS-2014. Tashkent-

2014. – P. 332-338.

24. Хўжаев О.К., Матякубов М.Я., Аллаберганов Д.Ю. Қарор қабул қилишга кўмаклашиш тизимлари ва маълумотларни интеллектуал таҳлили дастурий воситаларини яратишда дастурлаш тиллари ҳамда технологиялари таҳлили// «Муҳаммад ал-Хоразмий издошлари» Республика илмий-техникавий анжумани. Урганч-2018. -Б.375-377.

25. Хўжаев О.К., Кутлиев С.П., Аллаберганов Д.Ю. Сервисга йўналтилган архитектура асосида маълумотларни интеллектуал таҳлили учун веб сервисни лойиҳалаш// «Муҳаммад ал-Хоразмий издошлари» Республика илмий-техникавий анжумани. Урганч-2018. -Б. 377-379.

26. Хўжаев О.К., Хусинов Ш.О., Артиков М.Э., Матякубов М.Я. REST архитектура асосида маълумотларни интеллектуал таҳлили учун веб сервис ишлаб чиқиш// «Муҳаммад ал-Хоразмий издошлари» Республика илмий-техникавий анжумани. Урганч-2018. -Б.379-382.

27. Хўжаев О.К., Рўзибоев О.Б., Тажимов С.Ш. “CDSS”-клиникал қарор қабул қилишни қўллаб қувватлаш дастури// O‘zbekiston respublikasi intellektual mulk agentligi, Elektron hisoblash mashinalari uchun yaratilgan dasturning rasmiy ro‘yxatdan o‘tkazganligi to‘g‘risidagi guvohnoma №DGU04667. 11.08.2017

28. Хо‘jayев О.К., Artikov M.E., Ruzmetov A.X. “E-Cardio” – kardiologiya shifoxonalari uchun axborot va qaror qabul qilishga ko‘maklashish tizimi dasturiy ta‘minoti // O‘zbekiston respublikasi intellektual mulk agentligi, Elektron hisoblash mashinalari uchun yaratilgan dasturning rasmiy ro‘yxatdan o‘tkazganligi to‘g‘risidagi guvohnoma №DGU05637. 11.09.2018.

29. Хо‘jayев О.К., Artikov M.E., Ruzmetov A.X. “DM-API” intellectual taxlil dasturiy vositasi// O‘zbekiston respublikasi intellektual mulk agentligi, Elektron hisoblash mashinalari uchun yaratilgan dasturning rasmiy ro‘yxatdan o‘tkazganligi to‘g‘risidagi guvohnoma DGU05636. 11.09.2018

Автореферат «Муҳаммад ал-Хоразмий авлодлари» илмий журнали таҳририятида таҳрирландан ўтказилди ва ўзбек, рус ва инглиз тилларидаги матнларини мослиги текширилди.

Бичими 60x84¹/₁₆.Рақамли босма усули. Times гарнитураси.
Шартли босма табағи:3. Адади 85. Буюртма № 66.

Гувоҳнома reestr № 10-3719
«Тошкент кимё технология институти» босмахонасида чоп этилган.
Босмахона манзили: 100011, Тошкент ш., Навоий кўчаси, 32-уй.