

ТОШКЕНТ АХБОРОТ ТЕХНОЛОГИЯЛАРИ УНИВЕРСИТЕТИ
ҲУЗУРИДАГИ ИЛМий ДАРАЖАЛАР БЕРУВЧИ
DSc.13/30.12.2019.T.07.01 РАҚАМЛИ ИЛМий КЕНГАШ
АСОСИДАГИ БИР МАРТАЛИК ИЛМий КЕНГАШ

ТОШКЕНТ АХБОРОТ ТЕХНОЛОГИЯЛАРИ УНИВЕРСИТЕТИ

АКБАРАЛИЕВ БАХТИЁРЖОН БАКИРОВИЧ

РЕСУРС ЧЕКЛАНГАНДА ОБЪЕКТЛАРНИ ТАВСИФЛОВЧИ
ОПТИМАЛ ИНФОРМАТИВ ТИЗИМЛАРНИ ШАКЛЛАНТИРИШ
УСУЛИ ВА АЛГОРИТМЛАРИ

05.01.03 – Информатиканинг назарий асослари

ТЕХНИКА ФАНЛАРИ ДОКТОРИ (DSc) ДИССЕРТАЦИЯСИ
АВТОРЕФЕРАТИ

Тошкент – 2021

Докторлик (DSc) диссертацияси автореферати мундарижаси

Оглавление автореферата докторской (DSc) диссертации

Contents of the abstract Doctoral (DSc) dissertation

Акбаралиев Бахтиёржон Бакирович

Ресурс чекланганда объектларни тавсифловчи оптимал информатив тизимларни шакллантириш усул ва алгоритмлари..... 3

Акбаралиев Бахтиёржон Бакирович

Методы и алгоритмы формирования оптимальных информативных систем описания объектов в условиях ограниченности ресурсов 29

Akbaraliev Bakhtiyorjon Bakirovich

Methods and algorithms for the formation of optimal informative systems for describing objects in conditions of limited resources 55

Эълон қилинган ишлар рўйхати

Список опубликованных работ
List of published works 59

ТОШКЕНТ АХБОРОТ ТЕХНОЛОГИЯЛАРИ УНИВЕРСИТЕТИ
ҲУЗУРИДАГИ ИЛМий ДАРАЖАЛАР БЕРУВЧИ
DSc.13/30.12.2019.T.07.01 РАҚАМЛИ ИЛМий КЕНГАШ
АСОСИДАГИ БИР МАРТАЛИК ИЛМий КЕНГАШ

ТОШКЕНТ АХБОРОТ ТЕХНОЛОГИЯЛАРИ УНИВЕРСИТЕТИ

АКБАРАЛИЕВ БАХТИЁРЖОН БАКИРОВИЧ

РЕСУРС ЧЕКЛАНГАНДА ОБЪЕКТЛАРНИ ТАВСИФЛОВЧИ
ОПТИМАЛ ИНФОРМАТИВ ТИЗИМЛАРНИ ШАКЛЛАНТИРИШ
УСУЛИ ВА АЛГОРИТМЛАРИ

05.01.03 – Информатиканинг назарий асослари

ТЕХНИКА ФАНЛАРИ ДОКТОРИ (DSc) ДИССЕРТАЦИЯСИ
АВТОРЕФЕРАТИ

Тошкент – 2021

Техника фанлари бўйича фан доктори (DSc) диссертацияси мавзуси Ўзбекистон Республикаси Вазирлар Маҳкамаси ҳузуридаги Олий аттестация комиссиясида B2020.3.T370 рақам билан рўйхатга олинган.

Докторлик диссертацияси Тошкент ахборот технологиялари университетида бажарилган.

Диссертация автореферати уч тилда (ўзбек, рус, инглиз (резюме)) Илмий кенгашнинг веб-саҳифасида (www.tuit.uz) ва «ZiyoNet» Ахборот-таълим порталида (www.ziyounet.uz) жойлаштирилган.

Илмий маслаҳатчи:	Нишанов Аҳрам Хасанович техника фанлар доктори, доцент
Расмий оппонентлар:	Мухамедиева Дилноз Тулкуновна техника фанлари доктори, профессор Гулямов Шухрат Маннапович техника фанлари доктори, профессор Нурмухамедов Толаниддин Рамзиддинович техника фанлари доктори, доцент
Етакчи ташкилот:	Мирзо Улуғбек номидаги Ўзбекистон Миллий университети

Диссертация ҳимояси Тошкент ахборот технологиялари университети ҳузуридаги DSc.13/30.12.2019.T.07.01 рақамли Илмий кенгашнинг 2021 йил « 29 » апрель соат 14⁰⁰ даги мажлисида бўлиб ўтади. (Манзил: 100084, Тошкент шаҳри, Амир Темура 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 (199 рақам билан рўйхатга олинган). (Манзил: 100084, Тошкент, Амир Темура кўчаси, 108-уй. Тел.:(+99871) 238-65-44).

Диссертация автореферати 2021 йил « 17 » апрелда тарқатилди.
(2021 йил « 01 » апрелдаги 1 рақамли реестр баённомаси.)

Р.Х.Ҳамдамов

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

Ф.М.Нуралиев

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

М.А.Раҳматуллаев

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

КИРИШ (фан доктори (DSc) диссертацияси аннотацияси)

Диссертация мавзусининг долзарблиги ва зарурати. Дунёда катта ҳажмли маълумотларга ишлов беришда ўта муҳим параметрлардан бири бўлган ресурсларни ҳисобга олувчи маълумотларни интеллектуал таҳлил қилиш ва тимсолларни таниб олиш усул ва алгоритмларни ривожлантириш ва ишлаб чиқишга алоҳида эътибор берилмоқда. Объектларни аниқлаш учун зарур бўлган ресурсларни эътиборга олувчи таниб олиш, синфлаштириш ва объектларни тавсифловчи оптимал информатив тизимларни шакллантириш усул ва алгоритмларини ишлаб чиқиш ҳамда олинган назарий натижаларни амалий масалаларни ечишга тадбиқи бугунги куннинг ўта долзарб муаммоларидан ҳисобланади. Ҳозирги кунда таниб олишнинг мантиқий ва математик асосга эга усул ва алгоритмлари соҳа муаммоларига мосланганликлари билан ўта жозибадор бўлмоқда. Таниб олишнинг соҳа муаммоларига мослашувчан назарияси ва амалиётини ривожлантиришга нафақат Ўзбекистонда, хорижий, хусусан, Россия Федерацияси, Буюк Британия, Жанубий Корея, Япония, Сингапур, Ҳиндистон, АҚШ, Германия, Хитой ва бошқа мамлакатларда ҳам катта эътибор қаратилмоқда.

Жаҳонда ахборот коммуникация технологиялари, қурилмалар ва алоқа воситаларининг жадал суратлар билан ривожланиши, жумладан, турли хил ижтимоий тармоқларни кенг миқёсда тарқалиши натижасида инсоният тарафидан ишлаб чиқиладиган маълумотларнинг салмоғи ва тури кескин тарзда ортиб бормоқда. Бундан эса маълумотлар орасидан илгари номаълум бўлган ёки қўлланмаган, нотривиал, амалий жиҳатдан фойдали, талқин қилиш имконияти мавжуд, инсон фаолиятини турли соҳаларида қарор қабул қилиш учун зарур бўладиган маълумотларни ажратиб олиш, улар орасидаги яширин боғлиқлик ва қонуниятларни аниқлаш ҳамда тақдим этиладиган маълумотларни фойдаланувчи учун қулай ва тушунарли кўринишда ифодалаш зарурати юзага келмоқда. Бу, ўз навбатида, маълумотларни синфлаштириш, кластеризация қилиш, белгилар фазоси ўлчамини пасайтириш, объектларни тавсифловчи информатив тизимларни шакллантириш бўйича усул ва алгоритмларни ишлаб чиқишни талаб этади.

Сўнги йилларда мамлакатимизда ахборот коммуникация технологиялари, илмий-тадқиқот ва инновацион фаолиятни ривожлантириш ҳамда бу йўналишлардаги ютуқларни давлат ва жамият ҳаётидаги барча соҳаларга кенг жорий қилишга алоҳида эътибор берилмоқда. Буни тасдиғи сифатида 2017-2021 йилларда Ўзбекистон Республикасини ривожлантиришнинг бешта устувор йўналиши бўйича Ҳаракатлар стратегиясида аниқланган «...иқтисодиёт, ижтимоий соҳа, бошқарув тизимига ахборот-коммуникация технологияларини жорий этиш, ... илмий-тадқиқот ва инновация фаолиятини рағбатлантириш, илмий ва инновация ютуқларини амалиётга жорий этишнинг самарали механизмларини яратиш ...»¹ вазифалар хизмат қилади. Мазкур

¹ Ўзбекистон Республикаси Президентининг 2017 йил 7 февралдаги «Ўзбекистон Республикасини янада ривожлантириш бўйича Ҳаракатлар стратегияси тўғрисида» ПФ-4947-сон Фармони.

масалалар доирасида мамлакатимизда кўплаб илмий ва амалий изланишлар олиб борилмоқда. Бу йўналишда маълумотларни интеллектуал таҳлил қилишга, хусусан, ресурс чекланганда синфлаштириш, кластеризация ва объектларни тавсифловчи информатив тизимларни шакллантириш усул ва алгоритмларини ишлаб чиқишга ҳамда маълумотларни интеллектуал таҳлили учун дастурий таъминот яратишга алоҳида эътибор қаратилмоқда.

Ўзбекистон Республикасининг 2019 йил 29 октябрдаги «Илм-фан ва илмий фаолият тўғрисида»ги Қонуни, Ўзбекистон Республикаси Президентининг 2017 йил 7 февралдаги ПФ-4947-сон «Ўзбекистон Республикасини янада ривожлантириш бўйича ҳаракатлар стратегияси тўғрисида», 2018 йил 19 февралдаги ПФ-5349-сон «Ахборот технологиялари ва коммуникациялари соҳасини янада такомиллаштириш чора-тадбирлари тўғрисида», 2018 йил 21 сентябрдаги ПФ-5544-сон «2019-2021 йилларда Ўзбекистон Республикасини инновацион ривожлантириш стратегиясини тасдиқлаш тўғрисида»ги Фармонлари ва 2017 йил 15 мартдаги ПҚ-2834-сон «Тошкент ахборот технологиялари университетининг фаолиятини янада такомиллаштириш чора-тадбирлари тўғрисида» Қарорида ва бу соҳада қабул қилинган бошқа меъёрий-ҳуқуқий ҳужжатларда кўзда тутилган вазифаларни амалга оширишга ушбу диссертация тадқиқоти муайян даражада хизмат қилади.

Тадқиқотларнинг республика фан ва технологиялари ривожланишининг устувор йўналишларига мослиги. Ушбу тадқиқотлар фан ва технологиялари ривожланишининг IV «Ахборотлаштириш ва ахборот-коммуникацион технологияларни ривожлантириш» устувор йўналиши доирасида бажарилган.

Диссертация мавзуси бўйича хорижий илмий тадқиқотлар шарҳи².

Жаҳоннинг бир қатор етакчи илмий-тадқиқот марказлари ва олий таълим муассасаларидаги илмий мактаблар салмоқли маълумотларга ишлов бериш, маълумотларни интеллектуал таҳлил қилиш ва тимсолларни таниб олиш соҳаларида илмий фаолият олиб боришмоқда. Буларга мисол сифатида қуйидагиларни келтириб ўтиш мумкин: Massachusetts Institute of Technology, University of Florida, University of Illinois at Urbana-Champaign, IBM, Google, Honeywell International, Qualcomm ва Microsoft компаниялари (АҚШ), McGill University, Concordia University, Department of Electrical and Computer Engineering, University of Waterloo (Canada), Seoul National University, KAIST, Gachon University (Жанубий Корея), University of Chinese Academy of Sciences, Key laboratory of Big Data Mining and Knowledge Management, Chinese Academy of Sciences (Хитой), Nippon Electric Corporation (NEC) ва Hitachi компаниялари, University of Tokyo (Япония), Eindhoven University of Technology (Голландия), Anna University (Хиндистон), Czech Technical University in Prague (Чехия), National University of Singapore (Сингапур),

² Диссертация мавзуси бўйича илмий тадқиқотлар шарҳи <https://www.researchgate.net>, www.elsevier.com, <https://www.nxp.com>, <http://www.machinelearning.ru>, <http://www.bmstu.ru>, <https://www.msu.ru>, <https://www.bsuir.by>, <https://habr.com> ва бошқа манбалар асосида шакллантирилган.

Universiti Teknologi Malaysia, RMIT University (Австралия), Россия фанлар академияси Ҳисоблаш маркази, Н.Э.Бауман номидаги Москва Давлат техника университети, М.Ломоносов номидаги Москва Давлат университети, Новосибирск Давлат техника университети, Новосибирск Давлат университети, Санкт-Петербург техника университети (Россия Федерацияси), В.М.Глушков номидаги Кибернетика институти (Украина), Белоруссия давлат университети, Белоруссия Давлат информатика ва радиоэлектроника университети (Белоруссия) ва бошқалар.

Мазкур йўналишларда олиб борилган тадқиқотлар натижасида бир қатор натижалар олинган: маълумотларга дастлабки ишлов бериш, хусусан, маълумотлардаги «халақит» («шовқин»)ларни бартараф этиш, маълумотларни силлиқлаш, филтрлаш усул ва алгоритмлари яратилган (Massachusetts Institute of Technology, IBM, Google, Honeywell International, Microsoft, Россия фанлар академияси Ҳисоблаш маркази, Н.Э.Бауман номидаги Москва Давлат техника университети, М.Ломоносов номидаги Москва Давлат университети, Санкт-Петербург техника университети, В.М.Глушков номидаги Кибернетика институти); маълумотларни таниб олиш усул ва алгоритмлари ҳамда дастурий таъминотлар ишлаб чиқилган (Россия фанлар академияси Ҳисоблаш маркази, М.В.Ломоносов номидаги Москва Давлат университети, Massachusetts Institute of Technology, В.М.Глушков номидаги Кибернетика институти, Белоруссия Давлат информатика ва радиоэлектроника университети, Anna University, Czech Technical University in Prague, National University of Singapore, Gaziantep University, Universiti Teknologi Malaysia, RMIT University, IBM, Google, Honeywell International, Microsoft, Yandex); белгилар фазосининг ўлчамини камайтириш усуллари ва алгоритмлари ишлаб чиқилган (Seoul National University, KAIST, University of Chinese Academy of Sciences, Key laboratory of Big Data Mining and Knowledge Management, Chinese Academy of Sciences, NEC, University of Tokyo, Eindhoven University of Technology, Россия фанлар академияси Ҳисоблаш маркази, Н.Э.Бауман номидаги Москва Давлат техника университети, М.Ломоносов номидаги Москва Давлат университети, В.М.Глушков номидаги Кибернетика институти, Белоруссия давлат университети).

Жаҳонда маълумотларни интеллектуал таҳлил қилиш ва тимсолларни таниб олишнинг мавжуд усуллари ривожлантириш ва янгиларини яратиш бўйича қатор истиқболли йўналишларда тадқиқотлар олиб борилмоқда: катта хажмдаги маълумотларга дастлабки ишлов бериш; ресурсларни инобатга олган ҳолда тимсолларни таниб олиш, хусусан, объектларни тавсифловчи белгилар фазоси ўлчовини пасайтириш усул ва алгоритмларини ишлаб чиқиш; мавжуд усул, алгоритм ва дастурий таъминотларни турли соҳа муаммоларига мослаштириш.

Муаммонинг ўрганилганлик даражаси. Сўнги йилларда маълумотларнинг хажми ва маълумотларни тавсифловчи белгилар фазоси ўлчамини мунтазам ортиб бориши туфайли дунё олимлари томонидан маълумотларни интеллектуал таҳлили, хусусан, маълумотни (объектлар)

тавсифловчи оптимал информатив тизимларни ресурсга боғлиқ шакллантириш муаммоларига янада кўпроқ эътибор қаратилмоқда. Бунда тадқиқотлар икки хил йўналишда олиб борилмоқда: объектларни тавсифловчи бошланғич белгилар фазоси ўлчамини камайтириш, уларни оптималлаштириш ёки муҳим белгиларни қидириш назариясини ривожлантириш, яъни қатъий асосланган математик усул ва алгоритмларни ишлаб чиқиш; назарий тадқиқотларда олинган натижалар асосида амалий масалаларни ечишнинг алгоритмик ва дастурий таъминотларини ишлаб чиқиш. Маълумотларни интеллектуал таҳлил қилиш масалаларида синфлаштириш, кластеризация ва белгилар фазоси ўлчамини камайтириш хорижий олимлар М.А.Айзерман, В.В.Александров, М.Б.Айдарханов, Э.М.Браверман, В.Н.Вапник, В.И.Васильев, М.С.Ватанабе, А.Л.Горелик, Р.Гонсалес, Р.Дуда, П.М.Чеголин, Ю.И.Журавлев, Н.Г.Загоруйко, А.Г.Ивахненко, Г.С.Лбов, В.Д.Мазуров, Л.А.Растринг, М.И.Шлезингер ва бошқалар томонидан ишлаб чиқилган ва уларнинг ишларида баён қилинган.

Ўзбекистонда тимсолларни таниб олиш назария ва усулларини ривожланиши М.М.Камилов, тасвирларга рақамли ишлов бериш С.С.Садиқов, тимсолларни аниқлашда бўсақавий қийматлар назарияси Э.М.Алиев, тимсолларни таниб олиш назариясини тиббиётга тадбиқи Ф.Т.Адилова, баҳоларни ҳисоблаш алгоритмлари негизида информатив белгиларни танлаш масалалари тадқиқи З.Т.Адилова, турли типли маълумотларга дастлабки ишлов бериш назарияси Ш.Х.Фазилов, объектларга ишлов беришда дискрет дастурлаш назарияси ривожини Р.Х.Хамдамов, тимсолларни таниб олишда корректлик масалалари тадқиқи Н.А.Игнатъев, тимсолларни таниб олишда информативликнинг эвристик мезони асосида информатив белгилар мажмуасини танлаш назария ва амалиёти А.Х.Нишановларнинг номлари билан бевосита боғлиқ.

Мазкур соҳадаги тадқиқотлар таҳлили, шуни кўрсатадики, маълумотларга дастлабки ишлов бериш етарлича тадқиқ қилинмаган, хусусан, белгилар фазоси ўлчамини камайтириш ва синфлаштириш масалалари комплекс ҳолда қаралмаган ҳамда фойдаланилган ресурслар эътиборга олинмаган. Бундан ташқари кенг тарқалган информативлик мезони учун математик қатъий асосланган ҳамда ресурсни эътиборга олувчи усул ва алгоритмларни ишлаб чиқиш муаммоси етарли даражада ўрганилмаган.

Диссертация тадқиқотининг диссертация бажарилган олий таълим ёки илмий-тадқиқот муассасасининг илмий-тадқиқот ишлари режалари билан боғлиқлиги. Диссертация иши Муҳаммад ал-Хоразмий номидаги Тошкент ахборот технологиялари университети илмий тадқиқот режасининг №17–019 «Маълумотларни интеллектуал таҳлил қилиш тамойиллари асосида катта маълумотлар массивида ахборотларнинг янгилигини баҳолаш, тизимлаштириш ва талқин қилишнинг алгоритмик усуллари ва дастурий таъминотини ишлаб чиқиш» (2007-2011), А5-045 «Кенг кўламга мўлжалланган виртуал таълим тизими» (2012-2014), А5-051 «Видеоназорат тизимида динамик объектлар видеотасвирини қайта ишлаш алгоритмлари ва

дастурий таъминотини ишлаб чиқиш» (2015-2017), 22/19-Ф «Миллий электрон ҳукумат муҳити хизматларидан фойдаланиш самарадорлигини оширишнинг «iGov – маслаҳат-муҳокама-мониторинг тизими»» (2019-2020) лойиҳалари доирасида олиб борилган.

Тадқиқотнинг мақсади информативликнинг эвристик мезонлари негизида тадқиқот объектини аниқлашга ажратилган ресурсга боғлиқ оптимал информатив тизимларни шакллантириш усул ва алгоритмларини ишлаб чиқишдан иборат.

Тадқиқотнинг вазифалари:

диссертация тадқиқоти муаммоларини ўрганилганлик даражаларини аниқлаш мақсадида тимсолларни таниб олиш назариясининг ҳозирги ҳолати таҳлилини ўтказиш;

маълумотларни интеллектуал таҳлили, хусусан, синфлаштириш, кластеризация ва объектларни тавсифловчи информатив тизимларни шакллантириш масалаларини ҳал қилишда фойдаланиладиган информативлик мезонлар тадқиқи;

объектларни тавсифловчи оптимал информатив тизимларни шакллантириш усул ва алгоритмларини такомиллаштириш ва ишлаб чиқиш;

информатив тизимларни шакллантириш мезонига мос ҳал қилувчи қонидани қуриш;

синфлаштиришдаги ҳатоликни ҳисобга олувчи информатив тизимларни шакллантириш усул ва алгоритминини ишлаб чиқиш;

ресурс чекланганда объектларни тавсифловчи ℓ информатив тизимларни шакллантириш усул ва алгоритмларини ишлаб чиқиш;

синфлаштириш ва информатив тизимларни шакллантириш масалалари учун ўқув ва назорат танланмаларни шакллантириш йўллари аниқлаш;

ҳал қилувчи қонидани қуриш, ресурсга боғлиқ ва боғлиқсиз равишда объектларни тавсифловчи оптимал информатив тизимларни шакллантириш бўйича ишлаб чиқилган усулларни алгоритмик ва дастурий таъминотини яратиш ҳамда тажрибавий ва реал масалаларга тадбиқ қилиш.

Тадқиқотнинг объекти сифатида қийматлари «объект-хосса» ва «объект-хосса-синф» жадвали кўринишда тақдим этилган объектларнинг (жараёнлар, ходисалар, воқеалар) белгилари (тавсифлари, таснифлари, хусусиятлари, кўрсаткичлари) олинган.

Тадқиқотнинг предмети синфлаштириш, кластеризация ва информатив тизимларни шакллантиришда фойдаланиладиган мезонлар, ёндашувлар, усуллар ва алгоритмлардан иборат.

Тадқиқотнинг усуллари. Ишнинг назарий тадқиқотлари чизиқли ва чизиқсиз дастурлаш, оптималлаштириш, эҳтимоллар назарияси, математик таҳлил, дискрет математика, маълумотларни интеллектуал таҳлили ва тимсолларни аниқлаш усулларига асосланган, синфлаштириш ва ресурсга боғлиқ объектларни тавсифловчи информатив тизимларни шакллантириш ҳамда уларни ишончлилигини аниқлаш бўйича тажрибавий тадқиқотлар эса дастурий мажмуа орқали амалга оширилган.

Тадқиқотнинг илмий янгилиги қуйидагилардан иборат:

Фишер кўринишидаги информативлик мезонига нисбатан тартиблаш усули орқали шакллантирилган ℓ информатив тизимларнинг оптимал бўлишининг етарлилик шартлари аниқланган;

объектларни тавсифловчи белгилар турли турга тегишли бўлганда эвристик мезонлардан фойдаланган ҳолда белгилар орасидаги ўзаро боғлиқликни эътиборга олувчи оптимал ℓ информатив тизимларни шакллантиришнинг такомиллаштирилган усул ва алгоритми ишлаб чиқилган;

синфлаштиришнинг хатолик коэффицентига асосланган ҳолда синфлаштириш самарадорлик ва ишончлилиқ кўрсаткичини оширувчи, нотўғри синфлаштирилган объектларни аниқловчи ва таҳлил қилувчи оптимал ℓ информатив тизимларни шакллантириш усул ва алгоритми ишлаб чиқилган;

объектларни тавсифловчи оптимал ℓ информатив тизимларни шакллантириш мезони асосида синфлаштириш масаласида ҳал қилувчи қонидани қуриш усул ва алгоритми ишлаб чиқилган;

ихтиёрий турдаги информативлик мезонига нисбатан ресурсни эътиборга олган ҳолда объектларни тавсифловчи оптимал информатив тизимларни шакллантириш усул ва алгоритми ишлаб чиқилган;

информативликнинг Фишер типидagi эвристик мезонлари асосида объектларни аниқлашга ажратилган ресурс чекланганда объектларни тавсифловчи оптимал ℓ информатив тизимларни шакллантиришнинг самарали усул ва алгоритми ишлаб чиқилган;

маълумотларни интеллектуал таҳлили масалаларида танланмаларга қўйиладиган талаблар асосида тажрибавий ўқув ва назорат танланмаларни шакллантириш қонидаси такомиллаштирилган.

Тадқиқотнинг амалий натижалари қуйидагилардан иборат:

Фишер туридаги информативлик мезони асосида объектларни тавсифловчи бошланғич белгилар фазоси ўлчамини камайтириш имконини берувчи оптимал информатив тизимларни шакллантириш процедуралари ишлаб чиқилган;

синф ва объектларни муҳимлилиқ даражаларини инобатга олган ҳолда синфлаштиришдаги хатоликларни камайтириш ва нотўғри синфлаштирилган объектларни аниқлаш имконини берувчи процедура ишлаб чиқилган;

объектни аниқлаш учун ресурс чекланганда тўла танлов усули билан бир хил натижа тақдим қилувчи информатив тизимларни шакллантириш алгоритмлари яратилган;

синфлаштириш ва объектларни тавсифловчи информатив тизимларни шакллантириш бўйича таклиф этилаётган усул ва алгоритмларни ишончлилигини баҳолаш имконини берувчи ўқув ва назорат танланмаларини шакллантириш процедураси яратилган;

синфлаштириш, кластеризация ва объектларни тавсифловчи оптимал информатив тизимларни шакллантириш масалаларини ечишнинг мавжуд ва диссертацияда таклиф этилган самарали алгоритмларига мос сервисга

йўналтирилган архитектура асосида веб сервис типдаги дастурий мажмуа ишлаб чиқилган.

Тадқиқот натижаларининг ишончлилиги. Маълумотларни интеллектуал таҳлилида синфлаштириш, кластерлаш ва белгилар фазоси ўлчамини камайтириш масалаларини тўғри ва математик жиҳатдан коррект кўйилганлиги, шакллантирилган дискрет оптимизация масалаларининг ечимини топиш усуллариининг қатъий математик исботланган теорема, тасдиқ ва хоссалар кўринишида ифодаланганлиги, ҳисоблашлар, шунингдек ишлаб чиқилган усул ва алгоритмларнинг тажриба синовларида ҳамда модел ва реал масалаларда олинган натижаларни солиштирилганликлари билан изоҳланади.

Тадқиқот натижаларининг илмий ва амалий аҳамияти. Тадқиқот натижаларининг илмий аҳамияти маълумотларни интеллектуал таҳлили объектларини информатив тавсифини шакллантиришда эвристик информативлик мезонлари асосида белгилар фазоси ўлчамини камайтириш ва синфлаштиришнинг қатъий математик усуллариини ишлаб чиқиш орқали маълумотларни интеллектуал таҳлил қилиш, хусусан, тимсолларни таниб олиш назариясини ривожлантиришга хизмат қилади. Ишлаб чиқилган усул ва алгоритмлар ижтимоий соҳа, иқтисодиёт ва бошқарувда таниб олиш ва синфлаштириш тизимларини яратишда кенг кўламда қўлланилиши билан асосланади.

Тадқиқот натижаларининг амалий аҳамияти таклиф этилган алгоритмлар ва дастурий таъминот орқали турли замонавий таниб олиш тизимларини лойиҳалаш ва ишлаб чиқишда ишчи белгилар луғатини шакллантириш ва синфлаштириш билан изоҳланади. Тадқиқот натижаларини кўллаш объект, ҳодиса ва жараёнларнинг информатив тавсифи орқали идентификациялаш, таниб олиш ва синфлаштиришга сарфланадиган ресурсларни қисқартириш ва улардан оқилона фойдаланиш имконини беради.

Тадқиқот натижаларининг жорий қилиниши. Ресурс чекланганда объектларни тавсифловчи информатив тизимларни шакллантириш усул ва алгоритмлари бўйича олинган илмий натижалар асосида:

синфлаштиришнинг хатолик коэффицентига асосланган ҳолда синфлаштириш самарадорлик ва ишончлилиқ кўрсаткичини оширувчи, нотўғри синфлаштирилган объектларни аниқловчи ва таҳлил қилувчи оптимал ℓ информатив тизимларни шакллантириш усул ва алгоритми, информативликнинг Фишер типдаги эвристик мезонлари асосида объектларни аниқлашга ажратилган ресурс чекланганда объектларни тавсифловчи оптимал ℓ информатив тизимларни шакллантириш усул ва алгоритми ҳамда ҳал қилувчи қондани қуриш усул ва алгоритми асосида ишлаб чиқилган «ИВМТ_S» дастурий мажмуа Гидрогеология ва инженерлик геологияси институти ДК фаолиятига жорий қилинган (Ўзбекистон Республикаси Давлат Геология ва минерал ресурслар қўмитасининг 2020 йил 25 августдаги 03/07-сон маълумотномаси). Илмий тадқиқот натижалари гидрогеологик объектларни синфлаштириш ва кластеризация қилиш орқали

иш унумдорлигини ўртача 20-30% ошириш, турли сарф-харажатларни эса 10-15% камайтириш имконини берган;

объектларни тавсифловчи белгилар турли турга тегишли бўлганда эвристик мезонлардан фойдаланган ҳолда белгилар орасидаги ўзаро боғлиқликни эътиборга олувчи оптимал ℓ информатив тизимларни шакллантиришнинг такомиллаштирилган усул ва алгоритми, ва улар асосида ишлаб чиқилган «IBMT_S» дастурий мажмуа Ўзбекистон телекоммуникация тармоқларини бошқариш республика маркази ДУК фаолиятига жорий қилинган (Ахборот технологиялари ва коммуникацияларни ривожлантириш вазирлигининг 2020 йил 25 сентябрдаги 33-8/5595-сон маълумотномаси). Илмий тадқиқот натижалари телекоммуникация тармоқлари техник ҳолатини таснифловчи информатив параметрларни аниқлаш ҳамда телекоммуникация тармоқлари ва тармоқ фойдаланувчиларини синфлаштириш орқали меҳнат самарадорлигини ўртача 20-30% ошириш, вақт ва бошқа ресурслар сарфини 40% қисқартириш имконини берган;

ихтиёрий турдаги информативлик мезонига нисбатан ресурсни эътиборга олган ҳолда объектларни тавсифловчи оптимал информатив тизимларни шакллантириш усули ва улар асосида ишлаб чиқилган дастурий восита Ўзбекистон Республикаси Давлат активларини бошқариш агентлиги фаолиятига жорий қилинган (Ахборот технологиялари ва коммуникацияларни ривожлантириш вазирлигининг 2020 йил 25 сентябрдаги 33-8/5595-сон маълумотномаси). Илмий тадқиқот натижалари иш унумдорлигини ўртача 10-15% ошириш иш ҳажми ва молиявий харажатларни мос равишда 10-20% ва 5-8% га камайтириш имконини берган.

Тадқиқот натижаларининг апробацияси. Диссертациянинг асосий назарий ва амалий натижалари 11 та халқаро ва 5 та республика илмий-техник ва илмий-амалий анжуманларида маъруза ва муҳокама қилинган.

Тадқиқот натижаларининг эълон қилинганлиги. Тадқиқотлар мавзуси бўйича жами 35 та илмий иш эълон қилинган. Жумладан, Ўзбекистон Республикаси Олий аттестация комиссиясининг докторлик диссертациялари асосий илмий натижаларини чоп этиш тавсия этилган илмий нашрларда 15 та мақола, шулардан, 12 таси республика ва 3 таси хорижий журналларда нашр этилган, шунингдек ЭХМ учун яратилган дастурий маҳсулотларни рўйхатга олишда 4 та гувоҳнома олинган.

Диссертациянинг тузилиши ва ҳажми. Диссертация тузилиши кириш, бешта боб, хулоса, фойдаланилган адабиётлар рўйхати ва иловадан иборат. Диссертациянинг ҳажми 184 бетни ташкил этган.

ДИССЕРТАЦИЯ ИШИНING АСОСИЙ МАЗМУНИ

Кириш қисмида диссертация иши мавзусининг долзарблиги ва зарурияти асосланган, Ўзбекистон Республикаси фан ва технологиялари ривожланишининг устувор йўналишларига тадқиқотнинг мослиги кўрсатилган, тадқиқотнинг мақсад ва вазифалари белгилаб олиниб, тадқиқот объекти ва предмети аниқланган, олинган натижаларнинг ишончлилиги

асосланган, тадқиқотнинг илмий янгилиги, назарий ва амалий аҳамияти очиб берилган, тадқиқот натижаларининг амалиётга татбиқи, ишнинг синов натижалари, нашр этилган ишлар ва диссертациянинг тузилиши бўйича маълумотлар келтирилган.

Диссертациянинг биринчи «**Маълумотларни интеллектуал таҳлили масалаларини ҳал қилиш усул ва алгоритмлари**» бобида маълумотларни интеллектуал таҳлилининг асосий масалалари ва уларни ҳал қилиш босқичлари келтириб ўтилган, бундан ташқари синфлаштириш, кластеризация ва белгилар фазоси ўлчовини камайтириш бўйича мавжуд усуллар ва алгоритмлар назарий жиҳатдан таҳлил қилинган. Мазкур бобнинг сўнги параграфида олиб борилган таҳлил натижалари асосида диссертация тадқиқоти мақсади ва асосий масалалари ҳамда ушбу масалаларнинг математик қўйилиши шакллантирилган.

Фараз қилайлик, X ўқув танланма берилган бўлсин

$$\begin{aligned} X &= \bigcup_{p=1}^r X_p, X_p \cap X_q = \emptyset, p \neq q, \quad p, q = \overline{1, r}, \\ X_p &= \{x_{p1}, x_{p2}, \dots, x_{pm_p}\}, \end{aligned} \quad (1)$$

бунда $x_{pi} = (x_{pi}^1, x_{pi}^2, \dots, x_{pi}^N) \in D^N, i = \overline{1, m_p}, D^N$ – белгилар фазоси.

Танланма элементлари оддий ёки бир хил тасодифий тақсимланган (*independent identically distributed, i.i.d.*).

Умумийликка зид бўлмаган ҳолда белгилар фазосини белгилаш учун D^N ўрнига X дан фойдаланмиз.

$c = (c_1, c_2, \dots, c_N) \in R^N$ вектор орқали объектларни аниқлаш учун талаб этиладиган ресурсни белгилаб оламиз, бунда c_j ($j = \overline{1, N}$) объектнинг j – белгисини аниқлаш учун зарур бўлган ресурс харажатлари. Объект аниқлаш учун жами $c_0 \in R$ ресурс ажратилган бўлсин.

Изоҳ 1. Ресурс - бу бошланғич маълумотларни олиш учун сарфланадиган харажатлар миқдори.

Объектларни тавсифловчи оптимал информатив тизимларни шакллантириш (ОИТШ) да $\lambda = (\lambda^1, \lambda^2, \dots, \lambda^N)$ вектордан фойдаланамиз

$$\lambda: X \rightarrow X|_{\lambda} = \{x | x = (\lambda^1 x^1, \lambda^2 x^2, \dots, \lambda^N x^N)\}$$

бунда $\lambda^j \in \{0, 1\}$ ($j = \overline{1, N}$).

Агар $\lambda^j = 1$ бўлса, у ҳолда j -чи белги ОИТШда иштирок этади, акс ҳолда иштирок этмайди.

Диссертация ишида ҳал қилиниши кўзда тутилган масалаларнинг математик қўйилишини қуйидагича шакллантириб оламиз:

$$\text{I. } \begin{cases} I(\lambda) \rightarrow \text{extr} \\ \lambda \in \Lambda^{\ell} \end{cases}, \quad \text{II. } \begin{cases} I(\lambda) \rightarrow \text{extr} \\ \lambda \in \Lambda^{\ell} \\ (c, \lambda) \leq c_0 \end{cases},$$

бунда $I(\lambda)$ – информативлик мезони, I – ОИТШ (ресурсга боғлиқсиз), II – ОИТШ (ресурс чекланганда).

III. Шундай $\mathfrak{R}: X \rightarrow R$ функция (алгоритм) қуриш талаб этиладигани, мазкур функция учун қуйидаги ўринли бўлсин:

$\forall x \in X_p (p = \overline{1, r})$ учун $\mathfrak{R}(x, X_p |_\lambda) = \text{extr}_{q=\overline{1, r}} \mathfrak{R}(x, X_q |_\lambda)$,

бунда $X_p |_\lambda = \{x \in X_p | x = (\lambda^1 x^1, \lambda^2 x^2, \dots, \lambda^N x^N)\}$, x – синфлаштириш учун тақдим этилган объект.

IV. Ихтиёрий N, r натурал сонлар берилган бўлсин. У ҳолда шундай $X \subseteq R^N$ танланма шакллантириш талаб этиладики, у учун $X = \bigcup_{p=1}^r X_p$, $X_p \cap X_q = \emptyset, p \neq q, p, q = \overline{1, r}$ ўринли бўлсин.

Диссертациянинг иккинчи «**Маълумотларни интеллектуал таҳлили масалаларини ечишдаги информативлик мезонлар тадқиқи**» боби тимсолларни таниб олиш, синфлаштириш, кластеризация, муҳим белгиларни танлаб олиш масалаларини ҳал қилишда кенг фойдаланиладиган мезонлар таҳлиliga ҳамда муайян усул ва алгоритмларнинг самарадорлик, информативлик ва ишончлилик баҳоларига бағишланган.

2.1-параграфда асосий тушунчалар, белгилашлар ва таърифлар ҳамда ўқув танланма объектлари орасидаги яқинлик ёки ўхшашлик ўлчовларини аниқлаш йўллари келтириб ўтилган.

Фараз қилайлик, X ўқув танланма (1) кўринишида аниқланган бўлсин.

1. Агар $X \subset R^N$ бўлса, у ҳолда объектлар орасидаги яқинлик функцияси Эвклид метрикасидан фойдаланган ҳолда масофа кўринишида аниқланади.

2. Агар X танланма объектларини тавсифловчи белгилар $\{0,1\}$ ёки $\{1,2, \dots, K\}$ тўпламлардан қийматлар қабул қилса, у ҳолда $\forall x, y \in X$ объектлар орасидаги яқинлик функцияси (ρ) ва масофа (d) ни куйидагича аниқлаш мумкин:

$$a) \quad \rho(x^j, y^j) = \begin{cases} 1, & \text{агар } x^j = y^j \\ 0, & \text{агар } x^j \neq y^j \end{cases} \quad (2)$$

b) $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N$ – мусбат сонлар берилган бўлсин.

$$\rho(x^j, y^j) = \begin{cases} 1, & \text{агар } |x^j - y^j| \leq \varepsilon_j \\ 0, & \text{агар } |x^j - y^j| > \varepsilon_j \end{cases} \quad (3)$$

$$\rho(x, y) = \sum_{j=1}^N \rho(x^j, y^j) \quad (4)$$

$$d(x, y) = N - \rho(x, y) \quad (5)$$

Изоҳ 2. Агар белгилар $\{0,1\}$ ёки $\{1,2, \dots, K\}$ лардан қиймат қабул қилса, у ҳолда объектлар орасидаги яқинлик функцияси мос равишда (2) ёки (3) орқали аниқланади.

Таъриф 1. Агар $\sum_{j=1}^N \lambda^j = \ell$ бўлса, у ҳолда $\lambda = (\lambda^1, \lambda^2, \dots, \lambda^N)$ вектор ℓ информатив дейилади.

Барча ℓ информатив векторлар тўпламини Λ^ℓ орқали белгилаймиз:

$$\Lambda^\ell = \{ \lambda: \sum_{j=1}^N \lambda^j = \ell, \lambda^j \in \{0,1\}, j = \overline{1, N} \}.$$

Кўриниб турибдики, Λ^ℓ тўплам қуввати куйидагига тенг

$$|\Lambda^\ell| = C_N^\ell = \frac{N!}{\ell!(N-\ell)!}$$

У ҳолда

$$\Lambda^\ell: X \rightarrow X |_{\Lambda^\ell} = \{X |_\lambda | \lambda \in \Lambda^\ell\}. \quad (6)$$

Таъриф 2. $X|_{\lambda} \in X_{\Lambda^{\ell}}$ тизим $I(\lambda)$ га нисбатан оптимал дейилади, агар $\exists \lambda \in \Lambda^{\ell}$ учун $I(\lambda) = \underset{\mu \in \Lambda^{\ell}}{\text{extr}} I(\mu)$ ўринли бўлсин.

У ҳолда ОИТШ масаласини оптималлаштириш масаласига келтириш мумкин

$$\begin{cases} I(\lambda) \rightarrow \text{extr} \\ \lambda \in \Lambda^{\ell} \end{cases}, \quad (7)$$

бунда $I(\lambda)$ – информативлик (самарадорлик) мезони.

$\lambda \in \Lambda^{\ell}$ векторга нисбатан берилган X_p ($p = \overline{1, r}$) синф объектлари орасидаги яқинлик ўлчови сифатида қуйидаги катталиқдан фойланамиз:

$$X \subset R^N \text{ да: } S_p(\lambda) = S_p(X_p)|_{\lambda} = \sqrt{\frac{1}{m_p} \sum_{i=1}^{m_p} \sum_{j=1}^N \lambda^j (x_{pi}^j - \bar{x}_p^j)^2}.$$

$x^j \in \{0,1\}$ ёки $x^j \in \{1,2, \dots, K\}$ да:

$$S_p(\lambda) = S_p(X_p)|_{\lambda} = \sqrt{\frac{2}{m_p(m_p - 1)} \sum_{i=1}^{m_p} \sum_{k=i+1}^{m_p} d(x_{pi}, x_{pk})|_{\lambda}}.$$

$\lambda \in \Lambda^{\ell}$ векторга нисбатан берилган X_p ва X_q ($p, q = \overline{1, r}$) синфлар объектлари орасидаги яқинлик ўлчови:

$X \subset R^N$ да:

$$R_{pq}(\lambda) = R_{pq}|_{\lambda} = \sqrt{\sum_{j=1}^N \lambda^j (\bar{x}_p^j - \bar{x}_q^j)^2}.$$

$x^j \in \{0,1\}$ ёки $x^j \in \{1,2, \dots, K\}$ да:

$$R_{pq}(\lambda) = R_{pq}|_{\lambda} = \sqrt{\frac{1}{m_p m_q} \sum_{i=1}^{m_p} \sum_{k=1}^{m_q} d(x_{pi}, x_{qk})|_{\lambda}}.$$

бунда $d(x, y)|_{\lambda} = \sum_{j=1}^N \lambda^j - \rho(x, y)|_{\lambda}$, $\bar{x}_p - X_p$ синф ўрталаштирилган объекти.

ОИТШ учун информативлик мезони сифатида қуйидаги функционалдан фойдаланамиз:

$$I(\lambda) = \frac{\sum_{p=1}^{r-1} \sum_{q=p+1}^r R_{pq}^2(\lambda)}{\sum_{p=1}^r S_p^2(\lambda)}.$$

Маълум бир белгилашлар киритиш орқали мазкур функционални қуйидаги кўринишга келтириб олиш мумкин

$$I(\lambda) = \frac{(a, \lambda)}{(b, \lambda)} \quad (8)$$

бунда $a = (a_1, a_2, \dots, a_N) \in R^N$, $b = (b_1, b_2, \dots, b_N) \in R^N$.

Диссертациянинг учинчи «**Объектларни тавсифловчи оптимал информатив тизимларни шакллантириш ва унга мос ҳал қилувчи қондани қуриш усул ва алгоритмларини ишлаб чиқиш**» бобида оптимал информатив тизимларни шакллантириш усул ва уларни амалга ошириш

алгоритмлари ҳамда оптимал ℓ информатив тизимларни шакллантириш мезонига мос ҳал қилувчи қондани қуриш усули таклиф этилган.

3.1-параграфда «Тартиблаш» усули учун ОИТШ етарлилик шартлари келтириб ўтилган. Фишер мезонига нисбатан ОИТШ ва уни сифатини баҳолаш масаласини қуйидагича ифодалаб олиш мумкин:

$$\begin{cases} I(\lambda) = \frac{(a, \lambda)}{(b, \lambda)} \rightarrow \max \\ \lambda \in \Lambda^\ell \end{cases} \quad (9)$$

a ва b векторлар компоненталари учун қуйидаги нисбатларни ҳисоблаб оламиз:

$$\frac{a_1}{b_1}, \frac{a_2}{b_2}, \dots, \frac{a_N}{b_N} \quad (10)$$

Маълумки, тартиблаш усули орқали шакллантирилган информатив тизимлар ҳар доим ҳам оптимал бўлавермайди. Қуйида тартиблаш усули орқали шакллантирилган информатив тизимларнинг оптимал бўлишининг етарлилик шартлари келтириб ўтилган.

Фараз қилайлик, $a_i \geq 0, b_i > 0$ ($i = \overline{1, N}$) учун қуйидаги ўринли бўлсин:

$$\frac{a_1}{b_1} \geq \frac{a_2}{b_2} \geq \dots \geq \frac{a_N}{b_N}. \quad (11)$$

У ҳолда $\lambda_0 = (\underbrace{1, 1, \dots, 1}_\ell, \underbrace{0, 0, \dots, 0}_{N-\ell})$ деб белгилаб оламиз.

У ҳолда қуйидагилар ўринли

Теорема 1. Агар ихтиёрий $\ell, 1 < \ell < N - 1$ ва $\forall k > \ell$ да

$$b_k > b_\ell \geq b_{\ell-1} \geq \dots \geq b_2 \quad (12)$$

бўлса, у ҳолда λ_0 (9) масала учун оптимал ечим бўлади.

Теорема 1 даги (12) шартни енгиллаштириш имконини берувчи баъзи хусусий ҳолларни кўриб ўтамиз, айнан, қуйидаги шартларни бажарилишини талаб этамиз:

$$\frac{a_j}{b_j} - \frac{a_{j+1}}{b_{j+1}} = \alpha > 0, j = \overline{1, N-1} \quad (13)$$

Теорема 2. $\ell = 2$ бўлсин. Агар a, b векторлар учун (11),(13) ва

$$b_k \geq \frac{b_1 b_2}{b_1(k-1) + b_2(k-2)}, \quad k = \overline{3, N}$$

ўринли бўлса, у ҳолда λ_0 вектор оптимал ечим.

Теорема 3. $\ell = 3$ бўлсин. Агар a, b векторлар учун (11),(13) ва

$$b_3 \geq \frac{b_1 b_2}{2b_1 + b_2}, b_k \geq \frac{(2b_1 + b_2)b_3}{b_1(k-1) + b_2(k-2) + b_3(k-2)}, k = \overline{4, N}$$

ўринли бўлса, у ҳолда λ_0 вектор оптимал ечим.

Бу каби шартларни ℓ ($\ell < N$) ихтиёрий қийматлари учун олиш мумкин. Бунда шарт кўриниши ℓ нинг аниқ қийматига боғлиқ бўлади.

Фараз қилайлик, a, b векторлар компоненталари учун (11) бажарилсин, лекин (12) бажарилмасин. Қуйидагича сонлар кетма-кетлигини аниқлаймиз:

$$a'_1 = a_1, b'_1 = b_1, a'_j = \alpha_j a_j, b'_j = \alpha_j b_j \quad (j = \overline{2, N}), \quad (14)$$

бу ерда $\alpha_2 > \frac{b_1}{b_2}, \alpha_3 > \alpha_2 \frac{b_2}{b_3}, \dots, \alpha_\ell > \alpha_{\ell-1} \frac{b_{\ell-1}}{b_\ell}$ ва $\alpha_t > \alpha_\ell \frac{b_\ell}{b_t}, t = \overline{\ell+1, N}$.

У ҳолда $a' = (a'_1, a'_2, \dots, a'_N)$ ва $b' = (b'_1, b'_2, \dots, b'_N)$ янги вектор компоненталари учун (11) ва (12) шартларни иккаласи ҳам ўринли бўлади. Шу сабабли уларга теорема 1 ўринли бўлади. Юқорида келтирилганларга асосланиб, шуни тасдиқлаш мумкинки, (i, j) белгилар жуфтлиги учун $\frac{a_i}{b_i} = \frac{a_j}{b_j}$ шарт ўринли бўлишидан, информатив тизимларда уларнинг муҳимлик даражаси бир ҳил бўлиши келиб чиқмайди, агар $a_i \neq a_j$ ёки $b_i \neq b_j$ бўлса.

Теорема 4. Агар $2 \leq t \leq \ell < k \leq N$ га тегишли ихтиёрий t, k индекслар учун $b_t \leq b_k, a_k \leq a_t$ ўринли бўлса, у ҳолда λ_0 (9) учун оптимал ечим бўлади.

Теорема 5. Агар $2 \leq i \leq \ell < j \leq N$ га тегишли ва $\lambda \in \Lambda^\ell$ векторнинг нолдан фарқли ихтиёрий (i, j) индекслари жуфтлиги учун $b_j \leq b_i$ бўлсин. У ҳолда

а) агар $\frac{(a, \lambda_0)}{(b, \lambda_0)} \leq \frac{a_j - a_i}{b_j - b_i}$ бўлса, у ҳолда λ_0 (9) учун оптимал ечим бўлади;

б) агар $\frac{(a, \lambda_0)}{(b, \lambda_0)} > \frac{a_j - a_i}{b_j - b_i}$ бўлса, у ҳолда $I(\lambda) > I(\lambda_0)$ бўлади.

$2 \leq i \leq \ell < j \leq N$ учун қуйидагича белгилашлар киритиб оламиз

$$K_1 = \left\{ (i, j) : \frac{(a, \lambda_0)}{(b, \lambda_0)} \leq \frac{a_j - a_i}{b_j - b_i} \right\}, K_2 = \left\{ (i, j) : \frac{(a, \lambda_0)}{(b, \lambda_0)} > \frac{a_j - a_i}{b_j - b_i} \right\}.$$

$$\sum_{(i, j) \in K_1} (b_j - b_i) \left(\frac{(a, \lambda_0)}{(b, \lambda_0)} - \frac{a_j - a_i}{b_j - b_i} \right) \geq \sum_{(i, j) \in K_2} (b_j - b_i) \left(\frac{(a, \lambda_0)}{(b, \lambda_0)} - \frac{a_j - a_i}{b_j - b_i} \right). \quad (15)$$

Теорема 6. Агар ихтиёрий $\lambda \in \Lambda^\ell$ учун (15) тенгсизлик ўринли бўлса, у ҳолда λ_0 оптимал, акс ҳолда $I(\lambda) > I(\lambda_0)$ бўлади.

3.2-параграфда эвристик мезон сифатида (8) кўринишидаги Фишер функционалидан фойдаланган ҳолда ОИТШ учун такомиллаштирилган «Самара» усули ва уни амалга ошириш алгоритми таклиф этилган.

a ва b вектор компоненталари учун $a_i > 0, b_i > 0, i = \overline{1, N}$ деб олиш мумкин.

Агар белгини аниқлаш учун ҳисоблаш (ўлчаш)лар мураккаблиги ва/ёки талаб этиладиган ресурслар турлича бўлса, у ҳолда ОИТШ учун (9) масалани барча оптимал ечимларини аниқлаш муҳим ҳисобланади.

(9) масалани ҳал қилиш учун ёрдамчи вектор-функционал киритиб оламиз:

$$\begin{aligned} \varphi(\lambda) &= a(b, \lambda) - b(a, \lambda), \\ \varphi(\lambda) &= (\varphi_1(\lambda), \varphi_2(\lambda), \dots, \varphi_N(\lambda)) \in R^N. \end{aligned}$$

Ушбу вектор-функционал ОИТШ учун стратегияни аниқлашда муҳим ҳисобланади.

$\varphi(\lambda)$ вектор-функционал учун қуйидаги тасдиқ ўринли бўлади.

Тасдиқ 1. $\lambda, \mu \in \Lambda^\ell$ бўлсин. $I(\mu) \geq I(\lambda)$ учун $(\varphi(\lambda), \mu) \geq 0$ шартни бажарилиши зарур ва етарли.

$\varphi(\lambda) \in R^N$ бўлгани учун унинг компоненталарини тартиблаб олиш мумкин, яъни

$$\varphi_{j_1}(\lambda) \geq \varphi_{j_2}(\lambda) \geq \dots \geq \varphi_{j_N}(\lambda) \quad (16)$$

(16) дан фойдаланган ҳолда қуйидагича $\mu(\lambda) = (\mu_1, \mu_2, \dots, \mu_N)$ градиент вектор аниқлаб оламиз, бунда

$$\begin{aligned} \mu_{j_1}(\lambda) &= \mu_{j_2}(\lambda) = \dots = \mu_{j_\ell}(\lambda) = 1, \\ \mu_{j_{\ell+1}}(\lambda) &= \mu_{j_{\ell+2}}(\lambda) = \dots = \mu_{j_N}(\lambda) = 0. \end{aligned}$$

Кўриниб турибдики, $\mu(\lambda) \in \Lambda^\ell$ бўлиб, унинг учун қуйидаги ўринли

$$(\varphi(\lambda), \mu(\lambda)) = \max_{\eta \in \Lambda^\ell} (\varphi(\lambda), \eta).$$

Хосса 1. $\forall \lambda \in \Lambda^\ell$ ва $\mu(\lambda) \in \Lambda^\ell$ учун $(\varphi(\lambda), \mu(\lambda)) \geq 0$ ўринли.

Тасдиқ 2. Агар $I(\lambda) = I(\mu(\lambda))$ шарт ўринли бўлса, у ҳолда $\mu(\lambda)$ вектор (9) учун оптимал.

Умуман олганда, ҳар бир олдиндан берилган ℓ учун объектларни тавсифловчи оптимал ℓ информатив тизимлар бир нечта бўлиши мумкин, яъни (9) масала учун оптимал ечимлар бир нечта бўлиши мумкин.

Мисол. $N = 8$; $a = (6; 7; 5; 6; 4; 4,2; 4,1; 3)$ ва $b = (5; 6; 6; 7,2; 5; 5,5; 5,4; 4)$ бўлсин. У ҳолда $\ell = 4$ да $I(\lambda_1) = I(\lambda_2) = I(\lambda_3) = \max_{\lambda \in \Lambda^\ell} I(\lambda) = 1$ бўлади, бунда $\lambda_1 = (1,1,1,0,1,0,0,0)$; $\lambda_2 = (1,1,1,0,0,0,0,1)$; $\lambda_3 = (1,1,0,1,0,0,0,1)$.

Фараз қилайлик, бирор бир $\lambda \in \Lambda^\ell$ да $I(\lambda) = I(\mu(\lambda))$ шарт бажарилсин ва мазкур векторга мос $\varphi(\lambda)$ вектор компоненталари учун қуйидагилардан бири ўринли бўлсин:

$$\varphi_{j_1} = \varphi_{j_2} = \dots = \varphi_{j_N}, \quad (17)$$

$$\varphi_{j_1} \geq \varphi_{j_2} \geq \dots \geq \varphi_{j_\ell} > \varphi_{j_{\ell+1}} \geq \dots \geq \varphi_{j_N}, \quad (18)$$

$$\varphi_{j_1} \geq \dots \geq \varphi_{j_{\ell-t}} > \varphi_{j_{\ell-t+1}} = \dots = \varphi_{j_{\ell+l}} > \varphi_{j_{\ell+l+1}} \geq \dots \geq \varphi_{j_N}, \quad (19)$$

бунда t, l – мусбат бутун сонлар ва $\ell > t$, $\ell + l \leq N$.

Тасдиқ 3. Агар $\varphi(\lambda)$ компоненталари учун (17) шарт ўринли бўлса, у ҳолда $\forall \eta \in \Lambda^\ell$ (9) масала учун оптимал бўлади.

Демак, агар $\lambda \in \Lambda^\ell$ учун $I(\lambda) = I(\mu(\lambda))$ ва $\varphi(\lambda)$ учун (17) ўринли бўлса, у ҳолда (9) масала учун оптимал ечимлар сони C_N^ℓ га тенг бўлар экан.

Тасдиқ 4. Агар $\varphi(\lambda)$ компоненталари учун (18) шарт ўринли бўлса, у ҳолда $\mu(\lambda)$ (9) масаланинг ягона оптимал ечими бўлади.

Фараз қилайлик, $\varphi(\lambda)$ учун (19) ўринли бўлсин, у ҳолда бошланғич белгилар фазосини 3 та гуруҳга ажратиб оламиз, яъни

$$F_{gr} = \{j_1, j_2, \dots, j_{\ell-t}\}; S_{gr} = \{j_{\ell-t+1}, j_{\ell-t+2}, \dots, j_{\ell+l}\}; T_{gr} = \{j_{\ell+l+1}, j_{\ell+l+2}, \dots, j_N\}.$$

$$\Lambda^\ell(\mu(\lambda)) = \{\mu(\lambda) \in \Lambda^\ell: \mu^i = 1, i \in F_{gr}; \mu^j \in \{0,1\}, j \in S_{gr}; \mu^k = 0, k \in T_{gr}\}.$$

Бундан эса $\forall \mu(\lambda) \in \Lambda^\ell(\mu(\lambda))$ учун

$$(\varphi(\lambda), \mu(\lambda)) = \max_{\eta \in \Lambda^\ell} (\varphi(\lambda), \eta)$$

эканлиги келиб чиқади.

Тасдиқ 5. $\forall \zeta \in \Lambda^\ell(\mu(\lambda))$ вектор (9) масала учун оптимал ечим бўлади. Демак, бу ҳолда оптимал ечимлар сони $card(\Lambda^\ell(\mu(\lambda))) = C_{t+l}^t$ та бўлади.

3.3-параграфда объектларни тавсифловчи белгилар турли турга тегишли бўлган ҳолда синфлаштиришнинг хатолик коэффицентига асосланган оптимал информатив тизимларни шакллантириш усули таклиф этилган.

Фараз қилайлик, $\lambda \in \Lambda^\ell$ векторга нисбатан синфлаштиришдаги хатоликлар коэффицентини $\theta(\ell)|_\lambda$ ва нотўғри синфлаштирилган объектлар сони $\kappa(\ell)|_\lambda$ бўлсин. У ҳолда

$$\theta(\ell)|_\lambda = \frac{\kappa(\ell)|_\lambda}{m},$$

бу ерда $m = \sum_{p=1}^r m_p$ – ўқув танланмадаги объектларнинг умумий сони.

У ҳолда синфлаштиришнинг хатолик коэффицентига асосланган ОИТШ масаласини математик кўйилиши куйидагича бўлади:

Талаб этилади: $\ell \ll N$ учун шундай $\lambda \in \Lambda^\ell$ оптимал вектор аниқлансинки,
 $\theta(\ell) \leq \theta(N)$ ёки $\theta(\ell) \leq \theta(N) + \delta$

ўринли бўлсин, бунда δ рухсат этилган хатолик.

Таъриф 3. $X|_\lambda \in X|_{\Lambda^\ell}$ тизим ёки $\lambda \in \Lambda^\ell$ вектор оптимал дейилади, агар $\forall \mu \in \Lambda^\ell$ учун $\theta(\ell)|_\lambda \leq \theta(\ell)|_\mu$ ўринли бўлса.

Таъриф 4. Агар $\theta(\ell) = \min_{\lambda \in \Lambda^\ell} \theta(\ell)|_\lambda$ бўлса, у ҳолда $X|_{\Lambda^\ell}$ информатив тизимларда синфлаштириш хатолиги коэффицентини $\theta(\ell)$ га тенг дейилади.

Объектларни тавсифловчи ОИТШ учун куйидагича эҳтимоллик векторини киритиб оламиз:

$$p_\nu = (p_\nu^1, p_\nu^2, \dots, p_\nu^N),$$

бунда p_ν^j - j -чи белгини тизимга кириш эҳтимоллиги; ν – итерация тартиби.

$\nu = 1$ да $p_1 = (p_1^1, p_1^2, \dots, p_1^N)$ учун $p_1^j = \frac{1}{N}$; $j = \overline{1, N}$ деб оламиз.

Маълумки, агар $\lambda \in \Lambda^N$ ва $\mu \in \Lambda^k$ бўлса, у ҳолда $\lambda - \mu \in \Lambda^{N-k}$ бўлади.

Барча $\mu \in \Lambda^k$ векторлар учун $\theta(N-k)|_{\lambda-\mu}$ қиймат ҳисобланади ва куйидаги шарт текширилади

$$\theta(N-k)|_{\lambda-\mu} \leq \theta. \quad (20)$$

(20) шартни қаноатлантирувчи $\mu \in \Lambda^k$ векторга мос белгилар эҳтимолликлари қийматлари $\lambda - \mu$ векторга мос келувчи белгиларнинг эҳтимолликларига тенг тақсимлаб берилади, яъни

$$p_{\nu+1}(\mu) = p_\nu + \frac{(p_\nu, \mu)}{N-k} * p_\nu|_{\lambda-\mu} - p_\nu|_\mu, p_{\nu+1} = \frac{1}{T} \sum_\mu p_{\nu+1}(\mu),$$

бу ерда T - (20) шартни қаноатлантирувчи μ векторлар умумий сони.

Ҳар бир итерациядан сўнг $p_{\nu+1} = (p_{\nu+1}^1, p_{\nu+1}^2, \dots, p_{\nu+1}^N)$ эҳтимоллик вектори энг кичик k та компоненталарига мос белгилар белгилар фазосидан чиқариб юборилади ва N га янги қиймат ўзлаштирилади, яъни $N = N - k$.

Амалий масалалар ҳал қилинаётганда k учун $k \ll \ell$ шарт ўринли бўлиши мақсадга мувофиқ бўлиб, одатда, $k = 3 \div 5$ деб олиш тавсия этилади.

3.4-параграфда ОИТШ мезонига мос ҳал қилувчи қоидани куриш усули таклиф этилган.

Фараз қилайлик, $\{\mu_1, \mu_2, \dots, \mu_n\}$ - (1) кўринишида берилган X ўқув танланмага 3.2 параграфдаги таклиф этилган усул ва алгоритмни жорий

қилиш орқали олинган оптимал ℓ информатив векторлар тўплами бўлсин. У ҳолда ҳар бир μ_j ($j = \overline{1, n}$) векторга объектларни тавсифловчи бирор бир ℓ информатив тизим мос келади, яъни

$$\begin{aligned} \mu_j: X \rightarrow Y_{\mu_j}, \dim(X) = N, \dim(Y_{\mu_j}) = \ell, (j = \overline{1, n}), \\ \mu_j: X_p \rightarrow Y_p|_{\mu_j}, (p = \overline{1, r}). \end{aligned}$$

$\mathfrak{R}(y, Y_{\mu_j})$ орқали $Y_{\mu_j} = \bigcup_{p=1}^r Y_p|_{\mu_j}$, $Y_p|_{\mu_j} \cap Y_q|_{\mu_j} = \emptyset, p \neq q, p, q = \overline{1, r}$, $j = \overline{1, n}$ ўқув танланма учун ҳал қилувчи қоидани белгилаб оламиз, бунда y - синфлаштириш учун тақдим этилган объект.

ℓ ўлчовли Y_{μ_j} белгилар фазосида μ_j векторга мос Фишер мезони қуйидаги кўринишга эга бўлади:

$$I_{\mu_j} = \frac{\sum_{p=1}^{r-1} \sum_{q=p+1}^r R_{pq}^2|_{\mu_j}}{\sum_{p=1}^r S_p^2|_{\mu_j}} = \frac{\sum_{p=1}^{r-1} \sum_{q=p+1}^r \left\| \bar{y}_p|_{\mu_j} - \bar{y}_q|_{\mu_j} \right\|^2}{\sum_{p=1}^r \left(\frac{1}{m_p} \sum_{i=1}^{m_p} \left\| \bar{y}_p|_{\mu_j} - y_{pi}|_{\mu_j} \right\|^2 \right)}$$

$y \in Y_s|_{\mu_j}$ деган фарозда $\mathfrak{R}(y, Y_s|_{\mu_j})$ ҳал қилувчи қоида қуйидагича аниқланади:

$$\begin{aligned} \mathfrak{R}(y, Y_s|_{\mu_j}) &= I_{\mu_j}(y, Y_s|_{\mu_j}), \\ I_{\mu_j}(y, Y_s|_{\mu_j}) &= \frac{\sum_{\substack{p=1 \\ p \neq s}}^{r-1} \sum_{\substack{q=p+1 \\ q \neq s}}^r \left\| \bar{y}_p|_{\mu_j} - \bar{y}_q|_{\mu_j} \right\|^2 + \sum_{\substack{p=1 \\ p \neq s}}^{r-1} \left\| \bar{y}_p|_{\mu_j} - \bar{y}'_s|_{\mu_j} \right\|^2}{\sum_{\substack{p=1 \\ p \neq s}}^r \left(\frac{1}{m_p} \sum_{i=1}^{m_p} \left\| \bar{y}_p|_{\mu_j} - y_{pi}|_{\mu_j} \right\|^2 \right) + \frac{1}{m_s} \sum_{i=1}^{m_s} \left\| \bar{y}'_s|_{\mu_j} - y_{si}|_{\mu_j} \right\|^2}, \end{aligned}$$

бунда $\bar{y}'_s|_{\mu_j} = \frac{\bar{y}_s|_{\mu_j} + y}{2}$, $s = \overline{1, r}$, $\|*\|$ -масофа.

Агар $y \in Y_p|_{\mu_j}$ бўлса, у ҳолда $\mathfrak{R}(y, Y_p|_{\mu_j}) > \mathfrak{R}(y, Y_s|_{\mu_j})$ ($s = \overline{1, r}, s \neq p$).

ОИТШ информативлик мезонига мос ҳал қилувчи қоида қуйидагича амалга оширилади:

$$y \in Y_p|_{\mu_j} \Leftrightarrow \mathfrak{R}(y, Y_p|_{\mu_j}) = \max_{s=\overline{1, r}} \mathfrak{R}(y, Y_s|_{\mu_j}).$$

Фараз қилайлик, $Y_p|_{\mu_j}$ ($j = \overline{1, n}$) да $\mathfrak{R}(y, Y_{\mu_j})$ бўйича қуйидагилар аниқланган бўлсин

$$\kappa|_{\mathfrak{R}(y, Y_p|_{\mu_j})} = \alpha_p \cdot t_p^{out}|_{\mu_j} + \beta_p \cdot t_p^{in}|_{\mu_j}, \quad (21)$$

$$\theta|_{\mathfrak{R}(y, Y_{\mu_j})} = \frac{\sum_{p=1}^r (\alpha_p \cdot t_p^{out}|_{\mu_j} + \beta_p \cdot t_p^{in}|_{\mu_j})}{m} \quad (22)$$

бу ерда t_p^{out} ва t_p^{in} мос равишда $Y_p|_{\mu_j}$ синфдан хато чиқариб юборилган ва қўшиб олинган объектлар сони.

(21)-(22) ларни инобатга олган ҳолда оптимал ҳал қилувчи қоидани куйидаги масалаларни бирортасини ҳал қилиш орқали аниқлаб олиш мумкин:

$$\kappa|_{\mathfrak{R}(y, Y_{\mu^*})} = \min_{\mu_j} \kappa(\ell)|_{\mathfrak{R}(y, Y_{\mu_j})} \text{ ёки } \theta|_{\mathfrak{R}(y, Y_{\mu^*})} = \min_{\mu_j} \theta(\ell)|_{\mathfrak{R}(y, Y_{\mu_j})}.$$

Диссертациянинг тўртинчи «Ресурс чекланганда объектларни тавсифловчи ℓ информатив белгилар фазосини куриш усул ва алгоритмларини ишлаб чиқиш» бобида объектларни аниқлаш учун ресурс чекланганда оптимал информатив тизимларни шакллантириш усул ва алгоритмлари ҳамда Фишер туридаги информативлик мезони баъзи хоссалари келтириб ўтилган.

4.1-параграфда ресурс чекланганда ихтиёрий информативлик мезони учун ОИТШ усули келтириб ўтилган, яъни куйидаги масала ҳал қилинган

$$\begin{cases} I(\lambda) \rightarrow \text{extr} \\ \lambda \in \Lambda^\ell \\ (c, \lambda) \leq c_0 \end{cases}, \quad (23)$$

бунда $I(\lambda)$ – ихтиёрий информативлик мезони.

Таъриф 5. Агар $\lambda \in \Lambda^\ell$ ва $(c, \lambda) \leq c_0$ бўлса, у ҳолда λ вектор (23) масаланинг плани дейилади.

$\Lambda^\ell(c) = \{\lambda \in \Lambda^\ell: (c, \lambda) \leq c_0\}$ орқали барча планлар тўпламини белгилаб оламиз.

Фараз қилайлик, $c = (c_1, c_2, \dots, c_N)$ векторнинг барча компонентлари тартибланган бўлсин, яъни $\forall i (i = \overline{1, N-1})$ учун $c_i \leq c_{i+1}$. Мазкур вектор компоненталаридан фойдаланган ҳолда куйидагича кетма-кетликларни тузиб оламиз:

$$f_1 = \sum_{i=1}^{\ell} c_i, f_2 = \sum_{i=2}^{\ell+1} c_i, \dots, f_k = \sum_{i=k}^{\ell+k-1} c_i, k = N - \ell + 1.$$

Демак, $z \geq 2$ да $f_z = f_1 + \sum_{i=1}^{z-1} (c_{\ell+i} - c_i)$.

Тасдиқ 6. $\Lambda^\ell(c) \neq \emptyset$ фақат ва фақат $f_1 \leq c_0$ бўлса.

Энди эса $c_0 - f_k \geq 0$ шарт ўринли бўладиган k ларни аниқлаймиз.

Изох 3. Агар $f_{N-\ell+1} \leq c_0$ бўлса, у ҳолда $\Lambda^\ell(c) \equiv \Lambda^\ell$ бўлади.

Хосса 2. Агар $c_0 - f_1 \geq 0$ ва $c_0 - f_{N-\ell+1} < 0$ бўлса, у ҳолда $\exists p$ мавжудки $1 \leq p < N - \ell + 1$ куйидаги ўринли бўлади

$$\begin{cases} c_0 - f_p \geq 0 \\ c_0 - f_{p+1} < 0 \end{cases}. \quad (24)$$

Тасдиқ 7. Агар p (24) ечими бўлса, у ҳолда куйидаги ўринли бўлади $\Lambda_p^\ell = \{\lambda \in \Lambda^\ell: \lambda^j \in \{0,1\}, j = \overline{1, \ell+p-1}, \lambda^i \equiv 0, i = \overline{\ell+p, N}\} \subseteq \Lambda^\ell(c)$.

(c_1, c_2, \dots, c_N) векторда энг катта индексга эга ва ℓ информатив тизимларда иштирок этиши мумкин бўлган белги куйидаги тенгсизликлар системасини ҳал қилиш орқали аниқланади.

$$\begin{cases} f_0 + c_j \leq c_0 \\ f_0 + c_{j+1} > c_0 \end{cases}, \text{ бунда } f_0 = \sum_{i=1}^{\ell-1} c_i, j = \overline{\ell+p, N}. \quad (25)$$

Агар c_q ($q < N$) (25) нинг ечими бўлса, у ҳолда $c_{q+1}, c_{q+2}, \dots, c_N$ ларга мос келувчи белгилар белгилар фазосидан чиқарилади ва q ўлчовли белгилар фазосига ўтилади.

Демак, $\Lambda^\ell(q) = \{\lambda: \sum_{j=1}^q \lambda^j = \ell\}$. Бунга мос равишда $\Lambda^\ell(c)$ ва Λ_p^ℓ лар куйидаги кўринишга эга бўлади

$$\Lambda^\ell(c) = \{\lambda \in \Lambda^\ell(q): (c, \lambda) \leq c_0, c = (c_1, c_2, \dots, c_q)\},$$

$$\Lambda_p^\ell = \left\{ \lambda \in \Lambda^\ell(q): \sum_{j=1}^{\ell+p-1} \lambda^j = \ell, \sum_{j=\ell+p}^q \lambda^j = 0 \right\} \subseteq \Lambda^\ell(c).$$

Умуман олганда, ҳар бир f_k га ягона $\lambda \in \Lambda^\ell(q)$ мос келиб, мазкур вектор компоненталари учун куйидаги ўринли бўлади:

$$\lambda|_{f_k} = \left(\underbrace{0, 0, \dots, 0}_{k-1}, \underbrace{1, 1, \dots, 1}_\ell, \underbrace{0, 0, \dots, 0}_{q-(\ell+k-1)} \right).$$

У ҳолда $f_{q-\ell+1}$ га куйидаги вектор мос келади:

$$\lambda|_{f_{q-\ell+1}} = \left(\underbrace{0, 0, \dots, 0}_{q-\ell}, \underbrace{1, 1, \dots, 1}_\ell \right).$$

Куйида $\Lambda^\ell(c) \setminus \Lambda_p^\ell$ тўплагани $\lambda|_{f_k}$ лар орқали куриш йўлини келтириб ўтамай.

$\Lambda^\ell(c) \setminus \Lambda_p^\ell$ тўплагани $q - p$ нинг қийматидан келиб чиққан ҳолда куриш тавсия этилади, чунки, $q - p \leq 2\ell$ да мазкур тўплагани $\lambda|_{f_{q-\ell+1}}$ орқали куришда амаллар сони камроқ бўлади, аксинча, $q - p > 2\ell$ да эса $\lambda|_{f_k}$ ($k = \overline{1, p}$) ларга нисбатан куриш самарали ҳисобланади.

Демак, $\Lambda^\ell(c) \setminus \Lambda_p^\ell$ тўплагани куриш куйидагича амалга оширилади:

$q - p \leq 2\ell$ да $\Lambda_{f_{q-\ell+1}}^\ell(c)$ тўплагани, $\lambda|_{f_{q-\ell+1}}$ га нисбатан, куйидагича аниқланади:

$$\Lambda_{f_{q-\ell+1}}^\ell(c) = \{\lambda \in \Lambda^\ell(q): (c, \lambda) = (c, \lambda|_{f_{q-\ell+1}}) + \sum_{(i,j)} (c_i - c_j) \leq c_0, \\ i \in \{1, 2, \dots, \ell + p - 1\}, j \in \{\ell + p, \dots, q\}\} \subseteq \Lambda^\ell(c).$$

$q - p > 2\ell$ да $\Lambda_{f_k}^\ell(c)$ ($k = \overline{1, p}$) тўплагани, $\lambda|_{f_k}$ га нисбатан, куйидагича аниқланади:

$$\Lambda_{f_k}^\ell(c) = \{\lambda \in \Lambda^\ell(q) \setminus \Lambda_p^\ell: (c, \lambda) = (c, \lambda|_{f_k}) + \sum_{(i,j)} (c_j - c_i) \leq c_0, \\ i \in \{1, 2, \dots, \ell + k - 1\}, j \in \{\ell + k, \dots, q\}\} \subseteq \Lambda^\ell(c).$$

У ҳолда (23) масалага эквивалент масалаларга эга бўламиз:

а) $q - p \leq 2\ell$ да

$$\begin{cases} I_1(\lambda) \rightarrow \text{extr} \\ \lambda \in \Lambda_p^\ell \end{cases} \quad \text{ва} \quad \begin{cases} I_{f_{q-\ell+1}}(\lambda) \rightarrow \text{extr} \\ \lambda \in \Lambda_{f_{q-\ell+1}}^\ell(c) \end{cases}$$

б) $q - p > 2\ell$ да

$$\begin{cases} I_1(\lambda) \rightarrow \text{extr} \\ \lambda \in \Lambda_p^\ell \end{cases} \quad \text{ва} \quad \begin{cases} I_{f_k}(\lambda) \rightarrow \text{extr} \\ \lambda \in \Lambda_{f_k}^\ell(c) \end{cases}$$

Кўриниб турибдики, (23) масаланинг ечими қуйидагича аниқланади:
 $q - p \leq 2\ell$ да

$$I(\lambda) = \text{extr}\{I_1(\lambda), I_{f_{q-\ell+1}}(\lambda)\}$$

акс ҳолда

$$I(\lambda) = \text{extr}\left\{I_1(\lambda), \text{extr}_{f_k}\left(I_{f_k}(\lambda)\right)\right\}.$$

4.2-параграфда фишер кўринишидаги эвристик информативлик мезонидан фойдаланган ҳолда ресурс чекланганда ОИТШ усул ва алгоритми таклиф қилинган. Агар ОИТШда иштирок эта олмайдиган белгиларни белгилар фазосидан чиқариб юборилса, у ҳолда (23) қуйидаги кўринишга эга бўлади:

$$\begin{cases} I(\lambda) = \frac{(a, \lambda)}{(b, \lambda)} \rightarrow \max \\ \lambda \in \Lambda^\ell \\ (c, \lambda) \leq c_0 \end{cases}, \quad (26)$$

бунда $a, b, c \in R^q, a_j > 0, b_j > 0, c_j > 0, \lambda = (\lambda^1, \lambda^2, \dots, \lambda^q)$.

Қуйидагича предикат функция киритиб оламиз:

$$s(\lambda) = \begin{cases} -1, & c_0 - (c, \lambda) < 0 \\ 1, & c_0 - (c, \lambda) \geq 0 \end{cases} \text{ бунда } \lambda \in \Lambda^\ell.$$

У ҳолда (26) қуйидаги кўринишга эга бўлади:

$$\begin{cases} I^*(\lambda) = \frac{(a, \lambda)}{(b, \lambda)} * s(\lambda) \rightarrow \max \\ \lambda \in \Lambda^\ell \end{cases} \quad (27)$$

Тасдиқ 1 га асосан $(\varphi(\lambda), \mu) \geq 0 \Leftrightarrow \frac{(a, \mu)}{(b, \mu)} \geq \frac{(a, \lambda)}{(b, \lambda)}$.

Бундан эса қуйидаги ўринли эканлиги келиб чиқади

Хосса 2. Агар $(\varphi(\lambda), \mu) \geq 0$ бўлса, у ҳолда

- $s(\lambda) = -1, s(\mu) = -1$ учун $I^*(\mu) \leq I^*(\lambda) < 0$;
- $s(\lambda) = -1, s(\mu) = 1$ учун $I^*(\mu) \geq -I^*(\lambda) > 0 > I^*(\lambda)$;
- $s(\lambda) = 1, s(\mu) = -1$ учун $I^*(\mu) \leq -I^*(\lambda) < 0 < I^*(\lambda)$;
- $s(\lambda) = 1, s(\mu) = 1$ учун $I^*(\mu) \geq I^*(\lambda) > 0$.

Фараз қилайлик, $I^*(\mu) \geq I^*(\lambda)$ бўлсин. У ҳолда

$$\begin{aligned} \frac{(a, \mu)}{(b, \mu)} s(\mu) \geq \frac{(a, \lambda)}{(b, \lambda)} s(\lambda) &\Leftrightarrow (a, \mu)(b, \lambda)s(\mu) - (a, \lambda)(b, \mu)s(\lambda) \geq 0 \Leftrightarrow \\ &\Leftrightarrow (a(b, \lambda)s(\mu) - b(a, \lambda)s(\lambda), \mu) \geq 0. \end{aligned}$$

Бундан эса қуйидаги ўринли бўлади

Хосса 3. Агар $I^*(\mu) \geq I^*(\lambda)$ бўлса, у ҳолда

- $s(\lambda) = -1, s(\mu) = -1$ учун $(\varphi(\lambda), \mu) \leq 0$;
- $s(\lambda) = -1, s(\mu) = 1$ учун $(a(b, \lambda) + b(a, \lambda), \mu) \geq 0$;
- $s(\lambda) = 1, s(\mu) = -1$ учун $(a(b, \lambda) + b(a, \lambda), \mu) \leq 0$;
- $s(\lambda) = 1, s(\mu) = 1$ учун $(\varphi(\lambda), \mu) \geq 0$.

Келтирилган хоссалардан кўриниб турибдики, агар ихтиёрий $\lambda, \mu \in \Lambda^\ell$ учун $(\varphi(\lambda), \mu)$ га боғлиқсиз равишда, агар $s(\lambda) = -1, s(\mu) = 1$ бўлса, у ҳолда,

хар доим $I^*(\mu) > I^*(\lambda)$ ва аксинча, яъни $s(\lambda) = 1$, $s(\mu) = -1$ да хар доим $I^*(\mu) < I^*(\lambda)$ ўринли бўлади.

Изоҳ 4. Асосий мақсад - (27) учун оптимал векторларни топиш, бундай ҳолатда, берилган λ векторга нисбатан танлаб олинаётган навбатдаги вектор оптимал ёки йўқлигини аниқлаш лозим бўлади. Демак, μ вектор оптимал бўлиши учун, $s(\mu) = 1$ шарт бажарилиши талаб этилади.

Изоҳ 5. Берилган $\lambda \in \Lambda^\ell$ вектор (27) учун план бўлади, агар $s(\lambda) = 1$ бўлса.

$(\varphi(\lambda), \mu * s(\mu)) \geq 0$ ҳолатни таҳлил қиламиз.

$$\begin{aligned} (\varphi(\lambda), \mu * s(\mu)) &= (a(b, \lambda) - b(a, \lambda), \mu * s(\mu)) = (a, \mu)(b, \lambda)s(\mu) - \\ &- (b, \mu)(a, \lambda)s(\mu) \geq 0 \Leftrightarrow \frac{(a, \mu)}{(b, \mu)}s(\mu) \geq \frac{(a, \lambda)}{(b, \lambda)}s(\mu) \Leftrightarrow I^*(\mu) \geq I(\lambda) * s(\mu) \end{aligned}$$

Маълумки, $\forall \lambda \in \Lambda^\ell$ учун $I(\lambda) \geq I^*(\lambda)$. Демак,

$$I^*(\mu) \geq I(\lambda) * s(\mu) \Leftrightarrow I^*(\mu) \geq I^*(\lambda) * s(\mu).$$

Бундан эса куйидаги ўринли эканлиги келиб чиқади.

Тасдиқ 8. $\forall \lambda, \mu \in \Lambda^\ell$ учун $I^*(\mu) \geq I^*(\lambda) * s(\mu)$ фақат ва фақат $(\varphi(\lambda), \mu * s(\mu)) \geq 0$.

Фараз қилайлик, берилган $\lambda \in \Lambda^\ell$ вектор учун $s(\lambda) = 1$ бўлсин.

Куйидагича оператор киритиб оламиз

$$\mu: \Lambda^\ell \rightarrow \Lambda^\ell,$$

бунда

$$(\varphi(\lambda), \mu * s(\mu)) = \max_{\substack{\eta \in \Lambda^\ell \\ s(\eta)=1}} (\varphi(\lambda), \eta * s(\eta))$$

Мазкур шартни қаноатлантирувчи μ векторни $\mu(\lambda)$ деб белгилаб оламиз. У ҳолда берилган λ ва $\mu(\lambda)$ векторлар учун куйидаги ўринли бўлади.

Хосса 4. Ихтиёрий $\lambda \in \Lambda^\ell$ учун $(\varphi(\lambda), \mu(\lambda) * s(\mu)) \geq 0$, бунда $s(\lambda) = 1$.

Тасдиқ 9. Агар $I^*(\lambda) = I^*(\mu(\lambda))$ шарт ўринли бўлса, у ҳолда $\mu(\lambda)$ вектор (27) масала учун оптимал ечим бўлади.

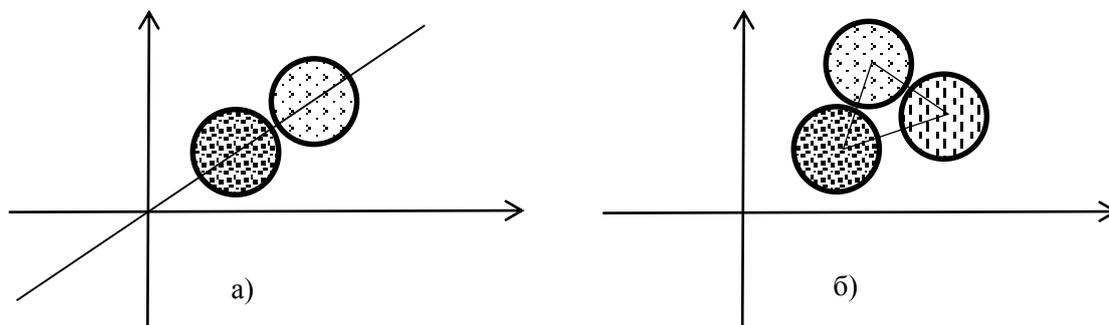
Умуман олганда, (27) масала учун оптимал ечимлар сони $\varphi(\lambda)$ вектор-компоненталари қийматларига боғлиқ бўлади.

4.3-параграфда ℓ информатив тизимларни шакллантиришда ҳамда ҳал қилувчи қондани қуришда фойдаланилган информативлик мезони, яъни Фишер кўринишидаги функционалнинг баъзи хоссалари келтириб ўтилган.

Диссертациянинг бешинчи «**Ишлаб чиқилган алгоритмларни модел ва амалий масалаларга тадбиқи тадқиқи**» бобида диссертация ишида таклиф этилган усул ва алгоритмларни амалий тадбиқи ҳамда ўқув ва назорат танланмаларини шакллантириш қондаси баён қилинган.

5.1-параграфда маълумотларни интеллектуал таҳлил қилиш бўйича мавжуд ва диссертацияда таклиф этилган усул ва алгоритмлар асосида яратилган «**IBMТ_S**» дастурлар мажмуасининг IDEF0 ва IDEF1 моделлари тавсифи келтириб ўтилган.

5.2-параграфда маълумотларни интеллектуал таҳлил қилиш модел масалалари учун гипотетик ўқув ва назорат танланмаларини шакллантириш қоидаси келтириб ўтилган. Мазкур қоида орқали ихтиёрий сондаги ва хажмдаги ўқув ва назорат танланма шакллантириш имконияти мавжудлиги асосланган.



1-расм. Синфларнинг гипотетик тимсоли: а) синфлар сони 2 та; б) синфлар сони 3 та.

5.3-параграфда юрак ишемик касалликлари (ЮИК) симптокомплексларини танлаш ва ташхислаш масаласи қаралган. Бунда ЮИКнинг уч кўриниши: X_1 – «Зўриқиш стенокардияси», X_2 – «Ўткир миокард инфаркти» ва X_3 – «Аритмик шакл» тадқиқ қилинган. Ҳар бир синф (кўриниш)да мос равишда 140, 120 ва 40 тадан объект (бемор) лар бўлиб, ҳар бир объект турли турга (миқдорий ва сифат) тегишли бўлган 62 та белги орқали тавсифланган.

Симптокомплексларни танлаш учун 3.2-параграфда, ташхислаш масаласи учун эса 3.4-параграфда таклиф этилган алгоритмлари тадбиқ қилинди. Натижада ташхислаш учун клиник белги ва симптомлар сонини 62 тадан 10 тагача камайтиришга эришилди. $\ell = 10$ да берилган функционалга экстремум қиймат берувчи 7 та симптокомплекслар аниқланди.

1-жадвал

ЮИК бўйича энг информатив симптокомплекслар

т/р	Танланган симптокомплекслар (ИБМ)
1.	$X_1, X_7, X_{16}, X_{22}, X_{27}, X_{36}, X_{41}, X_{51}, X_{57}, X_{62}$
2.	$X_1, X_7, X_{22}, X_{27}, X_{36}, X_{41}, X_{44}, X_{51}, X_{57}, X_{62}$
3.	$X_1, X_7, X_{22}, X_{27}, X_{36}, X_{41}, X_{47}, X_{51}, X_{57}, X_{62}$
4.	$X_1, X_7, X_{10}, X_{22}, X_{27}, X_{36}, X_{41}, X_{51}, X_{57}, X_{62}$
5.	$X_1, X_7, X_{22}, X_{27}, X_{34}, X_{36}, X_{41}, X_{51}, X_{57}, X_{62}$
6.	$X_1, X_2, X_7, X_{22}, X_{27}, X_{36}, X_{41}, X_{51}, X_{57}, X_{62}$
7.	$X_7, X_{22}, X_{27}, X_{34}, X_{36}, X_{41}, X_{47}, X_{51}, X_{57}, X_{62}$

5.4-параграфда ИВМТ_S дастурий мажмуани Фарғона вилояти худудида ер ости сувлари (ЕОС) гидродинамик ва гидрокимёвий кўрсаткичлари ўзгаришини тажрибавий тадқиқ қилиш учун жорий қилинганлиги натижалари келтириб ўтилган. Бунда қатор масалалар ҳал қилинган, хусусан, ЕОСни гидродинамик ва кидрокимёвий кўрсаткичлари ўзгариши тадқиқ қилинган; ЕОСнинг ифлослантирувчи таркибий қисмларининг асосий манбаларини

аниқлаш; муайян худудда жойлашган турли ифлослантивчи манбалар натижасида ифлосланган ЕОС ореол (чегара)ларини аниқлаш ва бошқалар.

Қуйида ер ости сув конларининг ифлосланиш даражаларини, уларнинг маъданли таркибига нисбатан аниқлаш натижалари келтириб ўтилган. Бунда ер ости сувлари шўрланганлик даражаси бўйича 6 та синфга ажратилган, яъни

$$X = \bigcup_{p=1}^6 X_p, X_p \cap X_q = \emptyset, p \neq q, p, q = \overline{1,6},$$

бунда X_1 –шўрланмаган, $m_1 = 749$; X_2 –кучсиз шўрланган, $m_2 = 294$; X_3 –ўртача шўрланган, $m_3 = 178$; X_4 –шўрланган, $m_4 = 77$; X_5 –кучли шўрланган, $m_5 = 151$; X_6 –ўта кучли шўрланган, $m_6 = 36$.

$x = (x^1, x^2, \dots, x^9) \in X$ -гидрогеологик объект, бунда x^1 – куруқ қолдиқ; x^2 – $НСО_3$; x^3 – $СО_3$; x^4 – NO_3 ; x^5 – Cl ; x^6 – SO_4 ; x^7 – Ca ; x^8 – Mg ; x^9 – $Na + K$.

Мазкур масалани «IBMТ_S» дастурий мажмуа ҳал қилиш орқали қуйидагиларга эга бўламиз.

2-жадвал

ОИТШ натижалари

ℓ	Ресурсга боғлиқсиз	$\max_{\lambda \in \Lambda^\ell} I(\lambda)$	Ресурсга боғлиқ	(c, λ)	$\max_{\lambda \in \Lambda^\ell} I(\lambda)$
4	1,2,3,6	4,3674	1,2,3,6	25	4,3674
5	1,2,3,4,6	4,1843	1,2,3,4,6	30	4,1843
6	1,2,3,4,5,6	4,0404	2,3,4,5,6,8	33	3,97
7	1,2,3,4,5,6,8	3,9513	ечим йўқ	$c > c_0$	

Танланган белгилар мажмуаси орқали амалга оширилган синфлаштириш натижалари қуйидаги жадвалда келтириб ўтилган.

3-жадвал

Синфлаштириш натижалари

ℓ	$\mathfrak{R}_{k-yq}(y, Y_p)$	$\mathfrak{R}_{BHA}(y, Y_p)$	$\mathfrak{R}_{XD}(y, Y_p)$	$\mathfrak{R}_{NT}(y, Y_p)$	$\mathfrak{R}_{Im}(y, Y_p)$
4	75,8	80,9	79,5	83,2	84,7
5	80	87,3	87,3	89,5	90
6	85,4	88,9	88	92	94

ХУЛОСА

«Ресурс чекланганда объектларни таснифловчи оптимал информатив тизимларни шакллантириш усул ва алгоритмлари» мавзусидаги докторлик (DSc) диссертацияси бўйича олиб борилган тадқиқотлар асосида қуйидаги хулосалар тақдим этилди:

1. Маълумотлар интеллектуал таҳлили, хусусан, синфлаштириш, кластеризация ва объектларни тавсифловчи информатив тизимларни шакллантириш масалаларини ҳал қилишда фойдаланиладиган усул ва алгоритмлар ҳамда информативлик мезонлари тадқиқ ва таҳлил қилинди. Бунинг натижасида, синфлаштириш, кластеризация ва белгилар фазоси ўлчамини камайтириш масалалари комплекс кўринишда ва ресурс чекланган шароитларда кам ўрганилганлиги аниқланди.

2. Маълумотларни интеллектуал таҳлили масалаларида ҳал қилувчи қонидани қуриш ва белгилар фазоси ўлчамларини камайтириш муаммолари формаллаштирилди. Бу ўз навбатида, ривожланган математик аппаратлардан унумли фойдаланган ҳолда ҳал қилувчи қонидани қуриш ва объектларни тавсифловчи информатив тизимларни шакллантириш бўйича қатъий асосланган усул ва алгоритмларни ишлаб чиқиш имконини берди.

3. Объектларни тавсифловчи оптимал информатив тизимларни шакллантиришда тартибланиш усули ҳар доим ҳам оптимал ечим тақдим этмаслиги асосланган. Фишер кўринишидаги информативлик мезонидан фойдаланган ҳолда тартибланиш усули учун шакллантирилган ℓ информатив тизимлар оптимал бўлишининг етарлилик шартлари аниқланди. Булар ўз навбатида, ушбу усулни қўлаш шартларини ва олинган натижаларнинг ишончлилигини аниқлашга имкон берди.

4. Информативликнинг эвристик мезонларидан фойдаланган ҳолда объектларни тавсифловчи оптимал ℓ информатив тизимларни шакллантиришнинг такомиллаштирилган усул ва алгоритми ишлаб чиқилди. Мазкур усул объектларни тавсифловчи белгилар турига боғлиқ эмас. Мазкур усул орқали олинган ечимлар оптимал ва оптимал ечимларни излаш процедураси 2-3 итерацияда яқунланиши исботланган. Оптимизация масаласи кўпгина критик нуқталарга эга бўлганда мазкур усул самарали эканлиги асосланган. Бу оптимал информатив тизимларни шакллантиришда ресурслардан самарали фойдаланишга хизмат қилади.

5. Синфлаштиришнинг хатолик коэффицентини эътиборга олувчи оптимал ℓ информатив тизимларни шакллантириш усул ва алгоритми ишлаб чиқилди. Бу синфлаштириш самарадорлиги ва ишончлилигини ошириш ҳамда нотўғри синфлаштирилган объектларни аниқлаш ва таҳлил қилиш имконини берди.

6. Объектларни тавсифловчи ℓ информатив тизимларни шакллантиришда фойдаланилган информативлик мезонига мос ҳал қилувчи қонидани қуриш усул ва алгоритми ишлаб чиқилди. Улар орқали синфлаштириш ва белгилар

фазосини камайтириш масалалари комплекс ҳолда ҳал қилинди. Булар синфлаштириш ишончилиги ва самарадорлигини оширишга хизмат қилади.

7. Объект белгиларини аниқлашга ажратилган ресурс чекланганлигини эътиборга олувчи ℓ информатив тизимларни шакллантириш усул ва алгоритми ишлаб чиқилган. Мазкур усулда, дастлаб, белгиларни аниқлашга ажратилган ресурс чекланганлигини эътиборга олувчи ℓ информатив λ векторлар $\Lambda^\ell(c) = \{\lambda \in \Lambda^\ell: (c, \lambda) \leq c_0\}$ тўплами куриб олинган. Мазкур ёндашувда белгилар тури ва информативлик мезони муҳим ҳисобланмайди. Бу фойдаланувчи арсеналида мавжуд ихтиёрий информативлик мезонидан фойдаланган ҳолда оптималлаштириш масаласини ҳал қилиш имконини беради.

8. Объект белгиларини аниқлашга ажратилган ресурс чекланганда Фишер типдаги информативлик мезони учун оптимал ℓ информатив тизимларни шакллантиришнинг самарали усул ва алгоритми таклиф қилинди. Белгиларни информатив тизимини шакллантиришнинг оптималлаштириш масаласини ҳал қилишда $I(\lambda)$ функционални энг тез ўсиш йўналишини кўрсатувчи $\varphi(\lambda)$ вектор-функциядан ва шакллантирилаётган информатив тизимларда белгилар мажмуасини иштирок этиш имкониятини кўрсатувчи $s(\lambda)$ предикат функциядан фойдаланилади. Белгилар турига нисбатан усулнинг универсаллиги кўрсатилган. Мазкур оптималлаштириш масаласининг ечими кам сонли белгилар ва кам харажатларда объектларни информатив тавсифлаш имконини берди.

9. Оптимал информатив тизимларни шакллантириш ва ҳал қилувчи қоидани куришда тузилмавий асос сифатида фойдаланилган Фишер туридаги информативлик мезоннинг хосса ва хусусиятлари ҳамда унинг самарадорлик баҳоси аниқланди. Тўла танлов усули орқали оптимал информатив тизимларни шакллантиришда таққослашлар сонини $C_{N-\ell}^\ell$ га камайтириш мумкинлиги кўрсатилди.

10. Маълумотларни интеллектуал таҳлил қилиш масалаларида танланмаларга қўйилган талаблар асосида модел ўқув ва назорат танланма шакллантириш қоидаси такомиллаштирилди. Бу объект белгиларини аниқлашга ажратилган ресурс чекланган ва чекланмаган шароитларда объектларни тавсифловчи информатив тизимларни шакллантириш ва синфлаштириш бўйича ишлаб чиқилган усул ва алгоритмлар ишончилиги ва самарадорлигини баҳолаш имкониятини берди.

11. Мавжуд ва диссертацияда таклиф этилган усул ва алгоритмлар асосида «IBMТ_S» дастурий мажмуа ишлаб чиқилди. Дастурий мажмуа иқтисодиёт, тиббиёт, телекоммуникация, гидрогеология ва қишлоқ хўжалиги соҳаларидаги амалий масалаларни ҳал қилишга жорий қилинган. Бу унумдорликни ўртача 20-30% га оширишга, сарф-харажатларни эса ўртача 10-15% камайтириш имкониятини берди.

**РАЗОВЫЙ НАУЧНЫЙ СОВЕТ ПРИ НАУЧНОМ СОВЕТЕ
DSc.13/30.12.2019.Т.07.01 ПО ПРИСУЖДЕНИЮ УЧЕНЫХ СТЕПЕНЕЙ
ПРИ ТАШКЕНТСКОМ УНИВЕРСИТЕТЕ ИНФОРМАЦИОННЫХ
ТЕХНОЛОГИЙ**

**ТАШКЕНТСКИЙ УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ
ТЕХНОЛОГИЙ**

АКБАРАЛИЕВ БАХТИЁРЖОН БАКИРОВИЧ

**МЕТОДЫ И АЛГОРИТМЫ ФОРМИРОВАНИЯ ОПТИМАЛЬНЫХ
ИНФОРМАТИВНЫХ СИСТЕМ ОПИСАНИЯ ОБЪЕКТОВ В
УСЛОВИЯХ ОГРАНИЧЕННОСТИ РЕСУРСОВ**

05.01.03 – Теоретические основы информатики

**АВТОРЕФЕРАТ
ДОКТОРСКОЙ (DSc) ДИССЕРТАЦИИ ПО ТЕХНИЧЕСКИМ НАУКАМ**

Ташкент – 2021

Тема докторской диссертации по техническим наукам (DSc) зарегистрирована в Высшей аттестационной комиссии при Кабинете Министров Республики Узбекистан за № В2020.3.Т370.

Диссертация выполнена в Ташкентском университете информационных технологий.

Автореферат диссертации на трех языках (узбекский, русский, английский (резюме)) размещен на веб-странице (www.tuit.uz) и на Информационно-образовательном портале «ZiyoNet» (www.ziyo.net).

Научный консультант:	Нишанов Ахрам Хасанович доктор технических наук, доцент
Официальные оппоненты:	Мухамедиева Дилноз Тулкуновна доктор технических наук, профессор Гулямов Шухрат Маннапович доктор технических наук, профессор Нурмухамедов Толаниддин Рамзиддинович доктор технических наук, доцент
Ведущая организация:	Национальный университет Узбекистана имени Мирзо Улугбека

Защита диссертации состоится « 29 » апреля 2021 года в 14⁰⁰ часов на заседании Научного совета 13/30.12.2019.Т.07.01 при Ташкентском университете информационных технологий. (Адрес: 100084, г. Ташкент, ул. Амира Темура, 108. Тел.: (99871) 238-64-43; факс: (99871) 238-65-52; e-mail: tuit@tuit.uz).

С диссертацией можно ознакомиться в Информационно-ресурсном центре Ташкентского университета информационных технологий (регистрационный номер № 199). (Адрес: 100084, г. Ташкент, ул. Амира Темура, 108. Тел.: (99871) 238-65-44).

Автореферат диссертации разослан « 17 » апреля 2021 года.
(реестр протокола рассылки № 1 от « 01 » апреля 2021 года.)

Р.Х.Хамдамов

Председатель научного совета по присуждению ученых степеней, д.т.н., профессор

Ф.М.Нуралиев

Ученый секретарь научного совета по присуждению ученых степеней, д.т.н., доцент

М.А.Рахматуллаев

Председатель научного семинара при Научном совете по присуждению ученых степеней, д.т.н., профессор

ВВЕДЕНИЕ (аннотация диссертации доктора наук (DSc))

Актуальность и востребованность темы диссертации. В мире особое внимание уделяется развитию и разработке методов и алгоритмов интеллектуального анализа данных и распознавания образов, учитывающих ресурсы, как важнейшие параметры при обработке больших данных. Одной из наиболее актуальных проблем сегодняшнего дня является разработка методов и алгоритмов распознавания, классификации и формирования оптимальных информативных систем описания объектов, учитывающих необходимые для определения объекта ресурсы, а также внедрение полученных теоретических результатов для решения практических задач. В настоящее время логические и математические методы и алгоритмы распознавания становятся все более привлекательными, поскольку они адаптируются к задачам данной области. Большое внимание уделяется развитию теории и практики распознавания, адаптированной к проблемам отрасли не только в Узбекистане, но и в зарубежных странах, в частности, в Российской Федерации, Великобритании, Южной Корее, Японии, Сингапуре, Индии, США, Германии, Китае и других странах.

В мире в результате бурного развития информационно-коммуникационных технологий, устройств и средств связи, в том числе, широкого распространения различных социальных сетей, резко растут объем и типы производимых человеком данных. Отсюда возникает необходимость извлечения из этих данных ранее неизвестной, нетривиальной, практически полезной и легко интерпретируемой информации, необходимой для принятия решений в различных сферах человеческой деятельности, обнаружение скрытой зависимости и закономерности, а также приведение представленной информации в удобный и понятный формат для пользователей. Это, в свою очередь, требует разработки методов и алгоритмов классификации, кластеризации, снижения размерности признакового пространства, а также формирования информативных систем описания объектов.

В последние годы в нашей стране особое внимание уделяется развитию информационно-коммуникационных технологий, научно-исследовательской и инновационной деятельности, а также широкому внедрению достижений в этом направлении во все отрасли экономики страны и социальную сферу. Подтверждением этого могут служить задачи, определенные в Стратегии действий по пяти приоритетным направлениям развития Республики Узбекистан на 2017-2021 годы «... внедрение информационно-коммуникационных технологий в экономику, социальную сферу, системы управления, ... стимулирование научно-исследовательской и инновационной деятельности, создание эффективных механизмов внедрения научных и инновационных достижений в практику...»³. В рамках решения этих задач в стране проводятся многочисленные научные и практические исследования. В

³ Указ Президента Республики Узбекистан №УП-4947 от 7 февраля 2017 г. «О Стратегии действий по дальнейшему развитию Республики Узбекистан»

этом направлении особое внимание уделяется интеллектуальному анализу данных, в частности, разработке методов и алгоритмов классификации, кластеризации, формированию оптимальных информативных систем описания объектов в условиях ограниченности ресурсов, а также созданию программного обеспечения для интеллектуального анализа данных.

Данное диссертационное исследование в определенной степени служит выполнению задач, предусмотренных Законом Республики Узбекистан от 29 октября 2019 года «О науке и научной деятельности», Указами Президента Республики Узбекистан №УП-4947 от 7 февраля 2017 г. «О Стратегии действий по дальнейшему развитию Республики Узбекистан», №УП-5349 от 19 февраля 2018 года «О мерах по дальнейшему совершенствованию сферы информационных технологий и коммуникаций», №ПФ-5544 от 21 сентября 2018 года «Об утверждении Стратегии инновационного развития Республики Узбекистан на 2019-2021 годы», №ПП-2158 от 3 апреля 2014 года «О мерах по дальнейшему внедрению информационно-коммуникационных технологий в реальном секторе экономики» и Постановлением Президента Республики Узбекистан №ПП-2834 от 15 марта 2017 года «О мерах по дальнейшему совершенствованию деятельности Ташкентского университета информационных технологий», а также другими нормативно-правовыми документами, принятыми в данной сфере.

Соответствие исследования приоритетным направлениям развития науки и технологий республики. Данное исследование выполнено в соответствии с приоритетным направлением развития науки и технологий Республики Узбекистан IV. «Развитие информатизации и информационно-коммуникационных технологий».

Обзор зарубежных научных исследований по теме диссертации⁴.

Ряд научных школ ведущих научно-исследовательских центров и высших учебных заведений мира ведут научную деятельность в области обработки больших данных, интеллектуального анализа данных и распознавания образов. В качестве примера можно привести следующие: Massachusetts Institute of Technology, University of Florida, University of Illinois at Urbana-Champaign, компании IBM, Google, Honeywell International, Qualcomm и Microsoft (США), McGill University, Concordia University, Department of Electrical and Computer Engineering, University of Waterloo (Canada), Seoul National University, KAIST, Gachon University (Южная Корея), University of Chinese Academy of Sciences, Key laboratory of Big Data Mining and Knowledge Management, Chinese Academy of Sciences (Китай), компании Nippon Electric Corporation (NEC) и Hitachi, University of Tokyo (Япония), Eindhoven University of Technology (Голландия), Anna University (Индия), Czech Technical University in Prague (Чехия), National University of Singapore (Сингапур), Universiti Teknologi Malaysia, RMIT University (Австралия), Вычислительный центр

⁴ Обзор научных исследований по теме диссертации составлен на основании <https://www.researchgate.net>, www.elsevier.com, <https://www.nxp.com>, <http://www.machinelearning.ru>, <http://www.bmstu.ru>, <https://www.msu.ru>, <https://www.bsuir.by>, <https://habr.com> и других источников

Российской академии наук, Московский государственный технический университет им. Н.Э.Баумана, Московский государственный университет им. М.Ломоносова, Новосибирский государственный технический университет, Новосибирский государственный университет, Санкт-Петербургский технический университет (Российская Федерация), Институт кибернетики им. В.М.Глушкова (Украина), Белорусский государственный университет, Белорусский государственный университет информатики и радиоэлектроники (Белоруссия) и другие.

В результате исследований в этих областях получен ряд результатов: созданы методы и алгоритмы предварительной обработки данных, в частности, устранения шумов в данных, сглаживания и фильтрации данных (Massachusetts Institute of Technology, IBM, Google, Honeywell International, Microsoft, Вычислительный центр Российской академии наук, Московский государственный технический университет им. Н.Э.Баумана, Московский государственный университет им. М.Ломоносова, Санкт-Петербургский технический университет, Институт кибернетики им. В.М.Глушкова); разработаны методы и алгоритмы распознавания образов, а также программное обеспечение (Вычислительный центр Российской академии наук, Московский государственный университет им. М.Ломоносова, Massachusetts Institute of Technology, Институт кибернетики им. В.М.Глушкова, Белорусский государственный университет информатики и радиоэлектроники, Anna University, Czech Technical University in Prague, National University of Singapore, Gaziantep University, Universiti Teknologi Malaysia, RMIT University, IBM, Google, Honeywell International, Microsoft, Yandex); разработаны методы и алгоритмы снижения размерности признакового пространства (Seoul National University, KAIST, University of Chinese Academy of Sciences, Key laboratory of Big Data Mining and Knowledge Managament, Chinese Academy of Sciences, NEC, University of Tokyo, Eindhoven University of Technology, Вычислительный центр Российской академии наук, Московский государственный технический университет им. Н.Э.Баумана, Московский государственный университет им. М.Ломоносова, Институт кибернетики им. В.М.Глушкова, Белорусский государственный университет).

Исследования по дальнейшему развитию существующих и созданию новых методов интеллектуального анализа данных и распознавания образов ведутся учеными всего мира по ряду перспективных направлений: разработка и усовершенствование методов предварительной обработки больших данных; разработка методов и алгоритмов распознавания образов, в частности, снижения размерности признакового пространства описывающих объектов с учетом ресурса; адаптация существующих методов, алгоритмов и программного обеспечения к различным отраслевым задачам.

Степень изученности проблемы. В последние годы, в связи с постоянным ростом объема данных и размерности признакового пространства, описывающих данных, учеными всего мира все больше внимания уделяется вопросам интеллектуального анализа данных, в

частности, формированию оптимальных информативных систем описания данных (объектов) с учетом ресурсов. Исследования ведутся в двух направлениях: развитие теоретических вопросов снижения размерности исходного признакового пространства, описывающего объекты, оптимизации и поиска важных признаков, т.е. разработка строго обоснованных математических методов и алгоритмов; разработка алгоритмического и программного обеспечения для решения конкретных практических задач на основе результатов теоретических исследований. Вопросы классификации, кластеризации и снижения размерности признакового пространства в задачах интеллектуального анализа данных, разработаны и изложены в работах таких зарубежных ученых, как М.А.Айзерман, В.В.Александров, М.Б.Айдарханов, Э.М.Браверман, В.Н.Вапник, В.И.Васильев, М.С.Ватанабе, А.Л.Горелик, Р.Гонсалес, Р.Дуд, П.М.Чеголин, Ю.И.Журавлев, Н.Г.Загоруйко, А.Г.Ивахненко, Г.С.Лбов, В.Д.Мазуров, Л.А.Растрингин, М.И.Шлезингер и другие.

В Узбекистане развитие теории и методов распознавания образов непосредственно связано с именем М.М.Камилова, цифровой обработки изображений – с именем С.С.Садыкова, теории пороговых значений при распознавании образов – с именем Э.М.Алиева, применение теории распознавания образов в медицине – с именем Ф.Т.Адыловой, исследование вопросов выбора информативных признаков на базе алгоритмов вычисления оценок – с именем З.Т.Адыловой, развитие теории предварительной обработки разнотипных данных – с именем Ш.Х. Фазылова, исследование вопросов корректности в распознавании образов – с именем Н.А. Игнатьева, развитие теории дискретного программирования при обработке данных – с именем Р.Х.Хамдамова, развитие теории и практики выбора набора информативных признаков на основе эвристической критериев информативности в распознавание образов – с именем А.Х. Нишановым.

Анализ исследований в этой области показывает, что недостаточно исследованы вопросы предварительной обработки данных, в частности, снижения размерности признакового пространства и классификации в комплексном виде, а также не учтены используемые при этом ресурсы. Кроме того, проблема разработки математически строго обоснованных и учитывающих ресурсы методов и алгоритмов для распространенных критериев информативности в достаточной степени не изучена.

Связь диссертационного исследования с планами научно-исследовательских работ высшего образовательного учреждения, где выполнена диссертация. Диссертационное исследование выполнено в рамках плана научно-исследовательских работ в Ташкентском университете информационных технологий №17-019 «Разработка методов алгоритмизации и программного обеспечения оценки новизны информации, структуризации и интерпретации в больших массивах данных на основе принципов интеллектуального анализа данных» (2007-2011), А5-045 «Виртуальная образовательная система широкого назначения» (2012-2014), №А5-051

«Разработка алгоритмов и программного обеспечения для обработки видеоизображений динамических объектов в системах видеонаблюдения» (2015-2017), 22/19-Ф «Разработка и внедрение «системы iGov - консалтинг-обсуждения-мониторинг» для повышения эффективности использования услуг национальной среды электронного правительства» (2019-2020).

Целью исследования является разработка методов и алгоритмов формирования оптимальных информативных систем с учетом выделенного ресурса для определения объекта на основе эвристических критериев информативности.

Задачи исследования:

провести анализ современного состояния теории распознавания образов с целью определения степени изученности проблемы диссертационного исследования;

исследование критериев информативности, используемых для решения задач интеллектуального анализа данных, в частности, задач классификации, кластеризации и формирования информативных систем описания объектов;

разработка и усовершенствование методов и алгоритмов формирования оптимальных информативных систем описания объектов;

построение решающего правила, адекватного критерию информативности формирования информативных систем;

разработка метода и алгоритма формирования информативных систем, учитывающие ошибки классификации;

разработка методов и алгоритмов формирования ℓ информативных систем описания объектов в условиях ограниченности ресурсов;

разработка правила генерации обучающей и контрольной выборки для задач классификации и формирования информативных систем;

создание алгоритмического и программного обеспечения предложенных методов построения решающего правила и формирования оптимальных информативных систем описания объектов с учетом ограниченности и неограниченности ресурсов, а также их применение при решении модельных и прикладных задач.

Объектом исследования являются признаки (характеристики, описание, свойства, показатели) объектов (процессов, явлений, событий), значения которых представлены в виде таблиц «объект-свойство» и «объект-свойство-класс».

Предметом исследования являются критерии, подходы, методы и алгоритмы, используемые при решении задач классификации, кластеризации и формировании информативных систем.

Методы исследования. Теоретические исследования в работе базировались на методах линейного и нелинейного программирования, оптимизации, теории вероятностей, математического анализа, дискретной математики, интеллектуального анализа данных и распознавания образов, а экспериментальные исследования по классификации и формированию

информативных систем описания объектов с учетом ресурсов и определения их надежности проводились с помощью программного комплекса.

Научная новизна диссертационного исследования заключается в следующем:

определены достаточные условия формирования оптимальности ℓ информативных систем для метода ранжировки относительно критерия информативности фишеровского типа;

разработаны усовершенствованный метод и алгоритм формирования оптимальных ℓ информативных систем с использованием эвристических критериев информативности, учитывающих взаимосвязь между признаками, где объекты описаны разнотипными признаками;

разработаны метод и алгоритм формирования оптимальных ℓ информативных систем, повышающих эффективность и надежность классификации, выявляющих и анализирующих неправильно классифицированные объекты на основе коэффициента ошибок классификации;

разработаны метод и алгоритм построения решающего правила в задаче классификации на основе соответствующего критерия информативности формирования оптимальных ℓ информативных систем;

разработаны метод и алгоритм формирования информативных систем описания объектов с учетом ресурсов относительно критерия информативности произвольного типа;

разработаны эффективный метод и алгоритм формирования оптимальных ℓ информативных систем при ограниченности ресурсов, выделенных для определения объекта, на основе эвристического критерия информативности фишеровского типа;

усовершенствовано правило генерации модельной обучающей и контрольной выборки на основе предъявляемых требований на выборки в задачах интеллектуального анализа данных.

Практические результаты исследования заключаются в следующем:

на основе критерия информативности фишеровского типа разработаны процедуры формирования оптимальных информативных систем позволяющие снижение размерности исходного признакового пространства описания объектов;

разработана процедура позволяющая снижение ошибки классификации и определение неправильно классифицированных объектов с учетом степени важности классов и объектов;

разработаны алгоритмы формирования информативных систем, результаты реализации которых совпадают с результатами метода полного перебора при ограниченности ресурсов для определения объектов;

разработана процедура формирования обучающей и контрольной выборки, позволяющая оценить надежность предложенных методов и алгоритмов классификации и формирование информативных систем описания объектов;

разработан программный комплекс типа веб-сервиса на основе сервисно-ориентированной архитектуры в соответствии с существующим и предложенным в диссертации эффективным алгоритмом для решения задач классификации, кластеризации и формирования информативных систем описания объектов.

Достоверность результатов исследования. Достоверность результатов исследования обосновывается корректностью постановки задач интеллектуального анализа данных классификации, кластеризации и снижения размерности признакового пространства, как задач дискретной оптимизации, строгостью математических выкладок, представленных в виде доказанных теорем, утверждений и свойств, а также сравнениями результатов экспериментальных исследований предложенных методов и алгоритмов на модельных и реальных задачах.

Научная и практическая значимость результатов исследования. Научная значимость результатов исследования служит дальнейшему развитию теории интеллектуального анализа данных, в частности, распознавания образов в части разработки математически строгих методов снижения размерности признакового пространства и классификации с использованием эвристических критериев информативности признаков при формировании информативного описания объектов интеллектуального анализа данных. Разработанные методы и алгоритмы обосновываются их широким применением при разработке систем распознавания и классификации в социальной сфере, экономике и управлении.

Практическая значимость результатов исследования определяется тем, что предложенные алгоритмы и программное обеспечение могут быть использованы при проектировании и разработке современных систем распознавания различного назначения в части формирования рабочего словаря признаков и классификации. Применение результатов исследования позволяет сократить и эффективно использовать ресурсы, требуемые для идентификации, распознавания и классификации, на основе информативного описания объектов, явлений и процессов.

Внедрение результатов исследования. На основе полученных научных результатов по методам и алгоритмам формирования информативных систем описания объектов в условиях ограниченности ресурса:

метод и алгоритм формирования оптимальных ℓ информативных систем, повышающий эффективность и надежность классификации, выявляющий и анализирующий неправильно классифицированные объекты на основе коэффициента ошибок классификации, метод и алгоритм формирования оптимальных ℓ информативных систем при ограниченности ресурсов, выделенных для определения объекта, на основе эвристического критерия информативности фишеровского типа, а также метод и алгоритм построения решающего правила, и разработанный на их основе программный комплекс «IBMT_S» внедрены в деятельность ГП Институт Гидрогеологии и инженерной геологии (справка Государственного комитета Геологии и

минеральных ресурсов Республики Узбекистан №03/07 от 25 августа 2020 г.). Результаты научного исследования позволили повысить производительность труда в среднем на 20-30%, уменьшить различные расходы на 10-15% за счет классификации и кластеризации гидрогеологических объектов;

усовершенствованный метод и алгоритм формирования оптимальных информативных систем с использованием эвристических критериев информативности, учитывающих взаимосвязь между признаками, где объекты описаны разнотипными признаками, и разработанный на их основе программный комплекс «IBMT_S» внедрены в деятельность ГУП Республиканский центр телекоммуникационных сетей Узбекистана (справка Министерства по развитию информационных технологий и коммуникаций №33-8/5595 от 25 сентября 2020 г.). Результаты научного исследования позволили повысить эффективность труда в среднем на 20-30%, сократить затраты времени и других ресурсов на 40% за счет определения информативных параметров, описывающих техническое состояние телекоммуникационных сетей, а также классификации сетей телекоммуникации и пользователей;

метод и алгоритм формирования информативных систем описания объектов с учетом ресурсов относительно критерия информативности произвольного типа, и разработанный на их основе программный комплекс «IBMT_S» внедрены в деятельность Агентства по управлению государственными активами Республики Узбекистан (справка Министерства по развитию информационных технологий и коммуникаций №33-8/5595 от 25 сентября 2020 г.). Результаты научного исследования позволили повысить производительность труда в среднем на 10-15%, снизить рабочую нагрузку и финансовые затраты на 10-20% и 5-8% соответственно.

Апробация результатов исследования. Теоретические и практические результаты диссертационного исследования докладывались и обсуждались на 11 международных и 5 республиканских научно-технических и научно-практических конференциях.

Опубликованность результатов исследования. По теме диссертации опубликовано 35 научных работ. Из них 15 научных статей, в том числе 12 в республиканских и 3 в зарубежных журналах, рекомендованных Высшей аттестационной комиссией Республики Узбекистан для публикации основных научных результатов докторских диссертаций, а также получены 4 свидетельства регистрации программных продуктов для ЭВМ.

Структура и объём диссертации. Диссертация состоит из введения, пяти глав, заключения, списка использованной литературы и приложений. Объём диссертации составляет 184 страница.

ОСНОВНОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ

Во введении обоснована актуальность темы диссертации в соответствии с приоритетными направлениями развития науки и технологий Республики Узбекистан, сформулированы цели и задачи, выявлены объекты и предметы

исследований, изложены научная новизна и практические результаты исследования, обоснована достоверность полученных результатов, раскрыта теоретическая и практическая значимости полученных результатов, приведены перечни внедрений результатов исследования в практику, сведения по опубликованным работам и структуре диссертации.

В первой главе диссертации «**Методы и алгоритмы решения задач интеллектуального анализа данных**» приведены основные задачи интеллектуального анализа данных и их этапы решения, кроме того, теоретически проанализированы существующие методы и алгоритмы классификации, кластеризации и снижения размерности признакового пространства. В последнем параграфе данной главы, на основе результатов проведенного анализа, сформулированы цель и основные задачи диссертационного исследования, а также математическая постановка этих задач.

Пусть дана обучающая выборка X

$$X = \bigcup_{p=1}^r X_p, X_p \cap X_q = \emptyset, p \neq q, \quad p, q = \overline{1, r}, X_p = \{x_{p1}, x_{p2}, \dots, x_{pm_p}\}, \quad (1)$$

где $x_{pi} = (x_{pi}^1, x_{pi}^2, \dots, x_{pi}^N) \in D^N, i = \overline{1, m_p}, D^N$ – признаковое пространство.

Элементы выборки являются простыми или независимыми одинаково распределенными (independent identically distributed, i.i.d.).

Не нарушая общности, для обозначения признакового пространства вместо D^N будем использовать X .

Обозначим через $c = (c_1, c_2, \dots, c_N) \in R^N$ – вектор ресурсов, необходимых для определения объекта, где c_j ($j = \overline{1, N}$) – ресурсные затраты на измерение j -го признака объекта. Пусть для определения объекта выделено всего ресурсов $c_0 \in R$.

Примечание 1. Ресурс – это величина расходов, связанных с получением первичной информации.

При формировании оптимальных информативных систем (ФОИС) описания объектов будем использовать вектор $\lambda = (\lambda^1, \lambda^2, \dots, \lambda^N)$, т.е.

$$\lambda: X \rightarrow X|_{\lambda} = \{x | x = (\lambda^1 x^1, \lambda^2 x^2, \dots, \lambda^N x^N)\},$$

где $\lambda^j \in \{0, 1\}$ ($j = \overline{1, N}$).

Если $\lambda^j = 1$, то j -й признак будет участвовать при ФОИС, в противном случае не будет участвовать.

Математическую постановку задач, решаемых в диссертации, можно сформулировать следующим образом:

$$\text{I. } \begin{cases} I(\lambda) \rightarrow \text{extr} \\ \lambda \in \Lambda^{\ell} \end{cases}, \quad \text{II. } \begin{cases} I(\lambda) \rightarrow \text{extr} \\ \lambda \in \Lambda^{\ell} \\ (c, \lambda) \leq c_0 \end{cases},$$

где $I(\lambda)$ – критерий информативности, I – ФОИС (без учета ресурса), II – ФОИС (с учетом ресурса).

III. Требуется построить функцию (алгоритм) $\mathfrak{R}: X \rightarrow R$, для этой функции выполнялось следующее:

$$\mathfrak{R}(x, X_p | \lambda) = \text{extr}_{q=\overline{1,r}} \mathfrak{R}(x, X_q | \lambda) \text{ для } \forall x \in X_p (p = \overline{1,r}),$$

здесь $X_p | \lambda = \{x \in X_p | x = (\lambda^1 x^1, \lambda^2 x^2, \dots, \lambda^N x^N)\}$, x – объект, предоставленный для классификации.

IV. Пусть даны N, r – произвольные натуральные числа. Требуется сформировать выборку $X \subseteq R^N$, для которого выполняется $X = \bigcup_{p=1}^r X_p$, $X_p \cap X_q = \emptyset, p \neq q, p, q = \overline{1,r}$.

Вторая глава диссертации «Исследование критериев информативности для решения задачи интеллектуального анализа данных» посвящена анализу широко используемых критериев информативности для решения задач распознавания образов, классификации, кластеризации и выбора важных признаков, а также оценки эффективности, информативности и надежности определенных методов и алгоритмов.

В параграфе 2.1 приведены основные понятия, обозначения и необходимые определения, а также способы определения меры близости или подобия между объектами обучающей выборки.

Допустим, обучающая выборка X определена в виде (1).

1. Если $X \subset R^N$, то функция близости между объектами определяется как расстояние с использованием Евклидовой метрики.

2. Если признаки объекта обучающей выборки X принимают значение из множества $\{0,1\}$ или $\{1,2, \dots, K\}$, то функция близости (ρ) и расстояние (d) между $\forall x, y \in X$ объектов определяются следующим образом:

$$\text{а) } \rho(x^j, y^j) = \begin{cases} 1, & \text{если } x^j = y^j \\ 0, & \text{если } x^j \neq y^j \end{cases} \quad (2)$$

б) пусть $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N$ – положительные числа.

$$\rho(x^j, y^j) = \begin{cases} 1, & \text{если } |x^j - y^j| \leq \varepsilon_j \\ 0, & \text{если } |x^j - y^j| > \varepsilon_j \end{cases} \quad (3)$$

$$\rho(x, y) = \sum_{j=1}^N \rho(x^j, y^j) \quad (4)$$

$$d(x, y) = N - \rho(x, y) \quad (5)$$

Примечание 2. Функция близости между объектами определяется из (2) или (3), если признаки принимает значение соответственно из $\{0,1\}$ или $\{1,2, \dots, K\}$.

Определение 1. Если $\sum_{j=1}^N \lambda^j = \ell$, то вектор $\lambda = (\lambda^1, \lambda^2, \dots, \lambda^N)$ называется ℓ информативным.

Множество всех ℓ информативных векторов обозначим через Λ^ℓ :

$$\Lambda^\ell = \{ \lambda: \sum_{j=1}^N \lambda^j = \ell, \lambda^j \in \{0,1\}, j = \overline{1,N} \}.$$

Отсюда видно, что мощность множества Λ^ℓ равна

$$|\Lambda^\ell| = C_N^\ell = \frac{N!}{\ell! (N - \ell)!}$$

Тогда

$$\Lambda^\ell: X \rightarrow X |_{\Lambda^\ell} = \{X | \lambda | \lambda \in \Lambda^\ell\}. \quad (6)$$

Определение 2. Назовем систему $X|_{\lambda} \in X_{\Lambda^{\ell}}$ оптимальным относительно $I(\lambda)$, если $\exists \lambda \in \Lambda^{\ell}$, для которого верно $I(\lambda) = \underset{\mu \in \Lambda^{\ell}}{\text{extr}} I(\mu)$.

Тогда задачу ФОИС можно свести к оптимизационной задаче

$$\begin{cases} I(\lambda) \rightarrow \text{extr} \\ \lambda \in \Lambda^{\ell} \end{cases}, \quad (7)$$

где $I(\lambda)$ – критерий информативности.

В качестве меры близости между всеми объектами заданного класса X_p ($p = \overline{1, r}$) относительно $\lambda \in \Lambda^{\ell}$ вектора будем использовать величину: при $X \subset R^N$:

$$S_p(\lambda) = S_p(X_p)|_{\lambda} = \sqrt{\frac{1}{m_p} \sum_{i=1}^{m_p} \sum_{j=1}^N \lambda^j (x_{pi}^j - \bar{x}_p^j)^2};$$

при $x^j \in \{0,1\}$ или $x^j \in \{1,2, \dots, K\}$:

$$S_p(\lambda) = S_p(X_p)|_{\lambda} = \sqrt{\frac{2}{m_p(m_p - 1)} \sum_{i=1}^{m_p} \sum_{k=i+1}^{m_p} d(x_{pi}, x_{pk})|_{\lambda}}.$$

В качестве меры близости между всеми объектами заданной пары X_p и X_q ($p, q = \overline{1, r}$) относительно $\lambda \in \Lambda^{\ell}$ вектора будем использовать величину: при $X \subset R^N$:

$$R_{pq}(\lambda) = R_{pq}|_{\lambda} = \sqrt{\sum_{j=1}^N \lambda^j (\bar{x}_p^j - \bar{x}_q^j)^2};$$

при $x^j \in \{0,1\}$ или $x^j \in \{1,2, \dots, K\}$:

$$R_{pq}(\lambda) = R_{pq}|_{\lambda} = \sqrt{\frac{1}{m_p m_q} \sum_{i=1}^{m_p} \sum_{k=1}^{m_q} d(x_{pi}, x_{qk})|_{\lambda}}.$$

Здесь

$$\bar{x}_p = \frac{1}{m_p} \sum_{i=1}^{m_p} x_{pi} = \left(\frac{\sum_{i=1}^{m_p} x_{pi}^1}{m_p}, \frac{\sum_{i=1}^{m_p} x_{pi}^2}{m_p}, \dots, \frac{\sum_{i=1}^{m_p} x_{pi}^N}{m_p} \right),$$

$$d(x, y)|_{\lambda} = N - \rho(x, y)|_{\lambda} \text{ или } d(x, y)|_{\lambda} = \sum_{j=1}^N \lambda^j - \rho(x, y)|_{\lambda}.$$

В качестве критерия информативности для ФОИС будем использовать следующий функционал:

$$I(\lambda) = \frac{\sum_{p=1}^{r-1} \sum_{q=p+1}^r R_{pq}^2(\lambda)}{\sum_{p=1}^r S_p^2(\lambda)}.$$

Введя некоторые обозначение, этот функционал можно свести к виду

$$I(\lambda) = \frac{(a, \lambda)}{(b, \lambda)} \quad (8)$$

где $a = (a_1, a_2, \dots, a_N) \in R^N$, $b = (b_1, b_2, \dots, b_N) \in R^N$.

В третьей главе диссертации «Разработка методов и алгоритмов формирования оптимальных информативных систем описания объектов и построения адекватного ему решающего правила» предложены методы и соответствующие им алгоритмы формирования оптимальных информативных систем, а также метод построения решающего правила, адекватного критерию формирования оптимальных ℓ информативных систем.

В параграфе 3.1 приведены достаточные условия ФОИС для метода ранжировки. Задачу ФОИС и оценки ее качества применительно к критерию Фишера можно выразить следующим образом:

$$\begin{cases} I(\lambda) = \frac{(a, \lambda)}{(b, \lambda)} \rightarrow \max, \\ \lambda \in \Lambda^\ell \end{cases} \quad (9)$$

Вычислим следующие отношения для компонент векторов a и b :

$$\frac{a_1}{b_1}, \frac{a_2}{b_2}, \dots, \frac{a_N}{b_N} \quad (10)$$

Как известно, информативные системы, сформированные методом ранжировки, не всегда оптимальны. Ниже приведены достаточные условия оптимальности информативных систем, формируемых методом ранжировки.

Пусть для $a_i \geq 0, b_i > 0$ ($i = \overline{1, N}$) верно следующее

$$\frac{a_1}{b_1} \geq \frac{a_2}{b_2} \geq \dots \geq \frac{a_N}{b_N}. \quad (11)$$

Тогда обозначим $\lambda_0 = (\underbrace{1, 1, \dots, 1}_\ell, \underbrace{0, 0, \dots, 0}_{N-\ell})$.

Справедлива следующая

Теорема 1. Если при любом ℓ , $1 < \ell < N - 1$ и $\forall k > \ell$

$$b_k > b_\ell \geq b_{\ell-1} \geq \dots \geq b_2, \quad (12)$$

то λ_0 является оптимальным решением для задачи (9).

Рассмотрим некоторые частные случаи, позволяющие ослабить условия (12) теоремы 1, а именно: потребуем выполнения следующих условий:

$$\frac{a_j}{b_j} - \frac{a_{j+1}}{b_{j+1}} = \alpha > 0, j = \overline{1, N-1} \quad (13)$$

Теорема 2. Пусть $\ell = 2$. Если для векторов a, b выполняются условия (11), (13) и справедливо

$$b_k \geq \frac{b_1 b_2}{b_1(k-1) + b_2(k-2)}, \quad k = \overline{3, N}.$$

Тогда λ_0 – оптимальное решение.

Теорема 3. Пусть $\ell = 3$. Если для векторов a, b выполняются условия (11), (13) и справедливо

$$b_3 \geq \frac{b_1 b_2}{2b_1 + b_2}, b_k \geq \frac{(2b_1 + b_2)b_3}{b_1(k-1) + b_2(k-2) + b_3(k-2)}, k = \overline{4, N}.$$

Тогда λ_0 – оптимальное решение.

Аналогичные условия можно получить и для произвольного значения ℓ ($\ell < N$). При этом вид условия зависит от конкретного значения ℓ .

Допустим для компонент векторов a и b выполняется условие (11), но не выполняется условие (12). Определим числовые последовательности следующим образом:

$$a'_1 = a_1, b'_1 = b_1, a'_j = \alpha_j a_j, b'_j = \alpha_j b_j \quad (j = \overline{2, N}), \quad (14)$$

где $\alpha_2 > \frac{b_1}{b_2}$, $\alpha_3 > \alpha_2 \frac{b_2}{b_3}$, ..., $\alpha_\ell > \alpha_{\ell-1} \frac{b_{\ell-1}}{b_\ell}$ и $\alpha_t > \alpha_\ell \frac{b_\ell}{b_t}$, $t = \overline{\ell+1, N}$.

Тогда для компонент новых векторов $a' = (a'_1, a'_2, \dots, a'_N)$ и $b' = (b'_1, b'_2, \dots, b'_N)$ выполняются условия (11) и (12). В связи с этим для них применима теорема 1. Исходя из вышеприведенного, можно утверждать, если для пар признаков (i, j) выполняется условие $\frac{a_i}{b_i} = \frac{a_j}{b_j}$ при $a_i \neq a_j$ (или $b_i \neq b_j$), то из этого не следует, что мера важности этих признаков одинакова в информативных системах.

Теорема 4. Если $b_t \leq b_k$, $a_k \leq a_t$ для любых t, k при $2 \leq t \leq \ell < k \leq N$, то λ_0 является оптимальным решением для (9).

Теорема 5. Пусть для индексов (i, j) ненулевых элементов вектора $\lambda \in \Lambda^\ell$ выполняются $b_j \leq b_i$, $2 \leq i \leq \ell < j \leq N$. Тогда

а) если $\frac{(a, \lambda_0)}{(b, \lambda_0)} \leq \frac{a_j - a_i}{b_j - b_i}$, то λ_0 является оптимальным решением для (7);

б) если $\frac{(a, \lambda_0)}{(b, \lambda_0)} > \frac{a_j - a_i}{b_j - b_i}$, то $I(\lambda) > I(\lambda_0)$.

Введем для $2 \leq i \leq \ell < j \leq N$ обозначения

$$K_1 = \left\{ (i, j) : \frac{(a, \lambda_0)}{(b, \lambda_0)} \leq \frac{a_j - a_i}{b_j - b_i} \right\}, K_2 = \left\{ (i, j) : \frac{(a, \lambda_0)}{(b, \lambda_0)} > \frac{a_j - a_i}{b_j - b_i} \right\}.$$

$$\sum_{(i, j) \in K_1} (b_j - b_i) \left(\frac{(a, \lambda_0)}{(b, \lambda_0)} - \frac{a_j - a_i}{b_j - b_i} \right) \geq \sum_{(i, j) \in K_2} (b_j - b_i) \left(\frac{(a, \lambda_0)}{(b, \lambda_0)} - \frac{a_j - a_i}{b_j - b_i} \right). \quad (15)$$

Теорема 6. Если для $\lambda \in \Lambda^\ell$ верно неравенство (15), то λ_0 является оптимальным, иначе $I(\lambda) > I(\lambda_0)$.

В параграфе 3.2 предлагается усовершенствованный метод и алгоритм «Эффект» для ФОИС с использованием эвристического критерия, заданного функционалом Фишера (8).

Можно полагать $a_i > 0$, $b_i > 0$, $i = \overline{1, N}$ для компонент векторов a и b .

Если сложность вычисления (измерения) и/или требуемый ресурс для определения признака различны, тогда для ФОИС важно найти все оптимальные решения задачи (9).

Для решения задачи (9) введем вспомогательный вектор-функционал:

$$\varphi(\lambda) = a(b, \lambda) - b(a, \lambda), \quad \varphi(\lambda) = (\varphi_1(\lambda), \varphi_2(\lambda), \dots, \varphi_N(\lambda)) \in R^N.$$

Данный вектор-функционал является существенным при определении стратегии для ФОИС.

Для вектор-функционала $\varphi(\lambda)$ справедливо следующее утверждение.

Утверждение 1. Пусть $\lambda, \mu \in \Lambda^\ell$. Для того, чтобы $I(\mu) \geq I(\lambda)$, необходимо и достаточно выполнение условия $(\varphi(\lambda), \mu) \geq 0$.

Так как $\varphi(\lambda) \in R^N$, то можно упорядочить ее компоненты, т.е.

$$\varphi_{j_1}(\lambda) \geq \varphi_{j_2}(\lambda) \geq \dots \geq \varphi_{j_N}(\lambda) \quad (16)$$

Используя (16) определим градиент-вектор $\mu(\lambda) = (\mu_1, \mu_2, \dots, \mu_N)$ следующим образом:

$$\mu_{j_1}(\lambda) = \mu_{j_2}(\lambda) = \dots = \mu_{j_\ell}(\lambda) = 1, \mu_{j_{\ell+1}}(\lambda) = \mu_{j_{\ell+2}}(\lambda) = \dots = \mu_{j_N}(\lambda) = 0.$$

Очевидно, что $\mu(\lambda) \in \Lambda^\ell$ и для него справедливо следующее

$$(\varphi(\lambda), \mu(\lambda)) = \max_{\eta \in \Lambda^\ell} (\varphi(\lambda), \eta).$$

Свойство 1. Для $\forall \lambda \in \Lambda^\ell$ и $\mu(\lambda) \in \Lambda^\ell$ справедливо условие $(\varphi(\lambda), \mu(\lambda)) \geq 0$.

Утверждение 2. Если $I(\lambda) = I(\mu(\lambda))$, то вектор $\mu(\lambda)$ является оптимальным для (9).

В общем, для каждого заранее заданного ℓ может существовать несколько оптимальных ℓ информативных систем описания объектов, т.е. могут существовать несколько оптимальных решений для задачи (9).

Пример. Пусть $N = 8$; $a = (6; 7; 5; 6; 4; 4,2; 4,1; 3)$ и $b = (5; 6; 6; 7,2; 5; 5,5; 5,4; 4)$. Тогда $I(\lambda_1) = I(\lambda_2) = I(\lambda_3) = \max_{\lambda \in \Lambda^\ell} I(\lambda) = 1$ при $\ell = 4$, где

$$\lambda_1 = (1,1,1,0,1,0,0,0); \lambda_2 = (1,1,1,0,0,0,0,1); \lambda_3 = (1,1,0,1,0,0,0,1).$$

Пусть для некоторого $\lambda \in \Lambda^\ell$ выполняется условие $I(\lambda) = I(\mu(\lambda))$ и справедливо одно из следующих соотношений для компонент вектора $\varphi(\lambda)$, соответствующий данному вектору:

$$\varphi_{j_1} = \varphi_{j_2} = \dots = \varphi_{j_N}, \quad (17)$$

$$\varphi_{j_1} \geq \varphi_{j_2} \geq \dots \geq \varphi_{j_\ell} > \varphi_{j_{\ell+1}} \geq \dots \geq \varphi_{j_N}, \quad (18)$$

$$\varphi_{j_1} \geq \dots \geq \varphi_{j_{\ell-t}} > \varphi_{j_{\ell-t+1}} = \dots = \varphi_{j_{\ell+l}} > \varphi_{j_{\ell+l+1}} \geq \dots \geq \varphi_{j_N}, \quad (19)$$

где t, l – целые положительные числа и $\ell > t, \ell + l \leq N$.

Утверждение 3. Если для компонент $\varphi(\lambda)$ выполняется условие (17), то $\forall \eta \in \Lambda^\ell$ является оптимальным решением для задачи (9).

Значит, если для $\lambda \in \Lambda^\ell$ верно $I(\lambda) = I(\mu(\lambda))$ и $\varphi(\lambda)$ верно (17), то количество оптимальных решений для задачи (9) равно C_N^ℓ .

Утверждение 4. Если для компонент $\varphi(\lambda)$ выполняется условие (18), то $\mu(\lambda)$ является единственным оптимальным решением для задачи (9).

Допустим, для $\varphi(\lambda)$ справедливо (19), тогда исходное признаковое пространство разделим на 3 группы, т.е.

$$F_{gr} = \{j_1, j_2, \dots, j_{\ell-t}\}; S_{gr} = \{j_{\ell-t+1}, j_{\ell-t+2}, \dots, j_{\ell+l}\}; T_{gr} = \{j_{\ell+l+1}, j_{\ell+l+2}, \dots, j_N\}.$$

$$\Lambda^\ell(\mu(\lambda)) = \{\mu(\lambda) \in \Lambda^\ell: \mu^i = 1, i \in F_{gr}; \mu^j \in \{0,1\}, j \in S_{gr}; \mu^k = 0, k \in T_{gr}\}.$$

Отсюда следует, что $\forall \mu(\lambda) \in \Lambda^\ell(\mu(\lambda))$ верно

$$(\varphi(\lambda), \mu(\lambda)) = \max_{\eta \in \Lambda^\ell} (\varphi(\lambda), \eta).$$

Утверждение 5. $\forall \zeta \in \Lambda^\ell(\mu(\lambda))$ вектор является оптимальным решением для задачи (9).

Значит, в данном случае количество оптимальных решений будет $\text{card}(\Lambda^\ell(\mu(\lambda))) = C_{t+l}^t$.

В параграфе 3.3 предложен метод формирования оптимальных информативных систем на основе коэффициента ошибки классификации, причем объекты описываются разнотипными признаками.

Пусть $\theta(\ell)|_\lambda$ – коэффициент ошибки классификации и $\kappa(\ell)|_\lambda$ – количество неправильно классифицированных объектов относительно $\lambda \in \Lambda^\ell$ вектора. Тогда

$$\theta(\ell)|_\lambda = \frac{\kappa(\ell)|_\lambda}{m},$$

где $m = \sum_{p=1}^r m_p$ – общее количество объектов обучающей выборки.

Тогда математическая постановка задачи ФОИС на основе коэффициента ошибки классификации выглядит следующим образом.

Требуется: При $\ell \ll N$ найти такой оптимальный вектор $\lambda \in \Lambda^\ell$, причем

$$\theta(\ell) \leq \theta(N) \text{ или } \theta(\ell) \leq \theta(N) + \delta,$$

где δ – допустимая погрешность.

Определение 3. Система $X|_\lambda \in X|_{\Lambda^\ell}$ или вектор $\lambda \in \Lambda^\ell$ называется оптимальным, если для $\forall \mu \in \Lambda^\ell$ верно $\theta(\ell)|_\lambda \leq \theta(\ell)|_\mu$.

Определение 4. Если $\theta(\ell) = \min_{\lambda \in \Lambda^\ell} \theta(\ell)|_\lambda$, то в информативных системах $X|_{\Lambda^\ell}$ коэффициент ошибки классификации считается равным $\theta(\ell)$.

Для ФОИС описания объектов введем вероятностный вектор:

$$p_\nu = (p_\nu^1, p_\nu^2, \dots, p_\nu^N),$$

где p_ν^j – вероятность вхождения j -ого признака в систему; ν – порядок итерации.

При $\nu = 1$ для $p_1 = (p_1^1, p_1^2, \dots, p_1^N)$ возьмем $p_1^j = \frac{1}{N}; j = \overline{1, N}$.

Известно, что если $\lambda \in \Lambda^N$ и $\mu \in \Lambda^k$, то $\lambda - \mu \in \Lambda^{N-k}$.

Для всех векторов $\mu \in \Lambda^k$ вычисляется значение $\theta(N-k)|_{\lambda-\mu}$ и проверяется условие

$$\theta(N-k)|_{\lambda-\mu} \leq \theta. \quad (20)$$

Значения вероятности признаков, соответствующих вектору $\mu \in \Lambda^k$, удовлетворяющих условию (20), равномерно распределяется между вероятностями признаков, соответствующих вектору $\lambda - \mu$, т.е.

$$p_{\nu+1}(\mu) = p_\nu + \frac{(p_\nu, \mu)}{N-k} * p_\nu|_{\lambda-\mu} - p_\nu|_\mu.$$

$$p_{\nu+1} = \frac{1}{T} \sum_{\mu} p_{\nu+1}(\mu),$$

где T – общее количество векторов μ , удовлетворяющих условию (20).

После каждой итерации удаляется k признаков, соответствующих наименьшим k компонентам $p_{v+1} = (p_{v+1}^1, p_{v+1}^2, \dots, p_{v+1}^N)$ вероятностного вектора и присваивается новое значение на N , т.е. $N = N - k$.

При решении практических задач целесообразно выполнение условия $k \ll \ell$ для k , и на практике обычно рекомендуется $k = 3 \div 5$.

В параграфе 3.4 предложен метод построения решающего правила, адекватного критерию ФОИС.

Пусть $\{\mu_1, \mu_2, \dots, \mu_n\}$ – множества оптимальных ℓ информативных векторов, полученных при применении метода и алгоритма, предложенных в параграфе 3.2, на обучающей выборке X , заданной в виде (1). Тогда каждому вектору μ_j ($j = \overline{1, n}$) соответствует некая ℓ информативная система описания объектов, т.е.

$$\begin{aligned} \mu_j: X &\rightarrow Y_{\mu_j}, \dim(X) = N, \dim(Y_{\mu_j}) = \ell, (j = \overline{1, n}), \\ \mu_j: X_p &\rightarrow Y_p|_{\mu_j}, (p = \overline{1, r}). \end{aligned}$$

Через $\mathfrak{R}(y, Y_{\mu_j})$ обозначим решающее правило для обучающей выборки $Y_{\mu_j} = \bigcup_{p=1}^r Y_p|_{\mu_j}$, $Y_p|_{\mu_j} \cap Y_q|_{\mu_j} = \emptyset, p \neq q, p, q = \overline{1, r}, j = \overline{1, n}$, где y – объект, предоставленный для классификации.

Критерий Фишера в ℓ -мерном признаковом пространстве Y_{μ_j} , соответствующий вектору μ_j , имеет следующий вид:

$$I_{\mu_j} = \frac{\sum_{p=1}^{r-1} \sum_{q=p+1}^r R_{pq}^2|_{\mu_j}}{\sum_{p=1}^r S_p^2|_{\mu_j}} = \frac{\sum_{p=1}^{r-1} \sum_{q=p+1}^r \left\| \bar{y}_p|_{\mu_j} - \bar{y}_q|_{\mu_j} \right\|^2}{\sum_{p=1}^r \left(\frac{1}{m_p} \sum_{i=1}^{m_p} \left\| \bar{y}_p|_{\mu_j} - y_{pi}|_{\mu_j} \right\|^2 \right)}$$

Решающее правило $\mathfrak{R}(y, Y_s|_{\mu_j})$ при предположении $y \in Y_s|_{\mu_j}$ определяется следующим образом:

$$\begin{aligned} \mathfrak{R}(y, Y_s|_{\mu_j}) &= I_{\mu_j}(y, Y_s|_{\mu_j}), \\ I_{\mu_j}(y, Y_s|_{\mu_j}) &= \frac{\sum_{\substack{p=1 \\ p \neq s}}^{r-1} \sum_{\substack{q=p+1 \\ q \neq s}}^r \left\| \bar{y}_p|_{\mu_j} - \bar{y}_q|_{\mu_j} \right\|^2 + \sum_{\substack{p=1 \\ p \neq s}}^{r-1} \left\| \bar{y}_p|_{\mu_j} - \bar{y}'_s|_{\mu_j} \right\|^2}{\sum_{\substack{p=1 \\ p \neq s}}^r \left(\frac{1}{m_p} \sum_{i=1}^{m_p} \left\| \bar{y}_p|_{\mu_j} - y_{pi}|_{\mu_j} \right\|^2 \right) + \frac{1}{m_s} \sum_{i=1}^{m_s} \left\| \bar{y}'_s|_{\mu_j} - y_{si}|_{\mu_j} \right\|^2}, \end{aligned}$$

где $\bar{y}'_s|_{\mu_j} = \frac{\bar{y}_s|_{\mu_j} + y}{2}$, $s = \overline{1, r}$, $\|*\|$ – расстояние.

Если $y \in Y_p|_{\mu_j}$, то $\mathfrak{R}(y, Y_p|_{\mu_j}) > R(y, Y_s|_{\mu_j})$ ($s = \overline{1, r}, s \neq p$).

Решающее правило, адекватное критерию информативности для ФОИС, реализуется следующим образом:

$$y \in Y_p|_{\mu_j} \Leftrightarrow \mathfrak{R}(y, Y_p|_{\mu_j}) = \max_{s=\overline{1, r}} \mathfrak{R}(y, Y_s|_{\mu_j}).$$

Допустим, в $Y_p|_{\mu_j}$ ($j = \overline{1, n}$) по $\mathfrak{R}(y, Y_{\mu_j})$ определены следующие

$$\kappa|_{\mathfrak{R}(y, Y_{\mu_j})} = \alpha_p \cdot t_p^{out}|_{\mu_j} + \beta_p \cdot t_p^{in}|_{\mu_j}, \quad (21)$$

$$\theta|_{\mathfrak{R}(y, Y_{\mu_j})} = \frac{\sum_{p=1}^r (\alpha_p \cdot t_p^{out}|_{\mu_j} + \beta_p \cdot t_p^{in}|_{\mu_j})}{m}. \quad (22)$$

Здесь t_p^{out} и t_p^{in} – количество неправильно исключенных и добавленных объектов для класса $Y_p|_{\mu_j}$ соответственно.

Учитывая (21)-(22), оптимальное решающее правило можно определить, решая одну из следующих задач:

$$\kappa|_{\mathfrak{R}(y, Y_{\mu^*})} = \min_{\mu_j} \kappa(\ell)|_{\mathfrak{R}(y, Y_{\mu_j})} \text{ или } \theta|_{\mathfrak{R}(y, Y_{\mu^*})} = \min_{\mu_j} \theta(\ell)|_{\mathfrak{R}(y, Y_{\mu_j})}.$$

В четвертой главе диссертации «**Разработка методов и алгоритмов формирования ℓ информативных систем описания объектов в условиях ограниченности ресурсов**» приведены методы и алгоритмы формирования оптимальных информативных систем описания объектов в условиях ограниченности ресурсов, а также некоторые свойства критерия информативности фишеровского типа.

В параграфе 4.1 приведен метод ФОИС для произвольного критерия информативности в условиях ограниченности ресурсов, т.е. решена следующая задача

$$\begin{cases} I(\lambda) \rightarrow extr \\ \lambda \in \Lambda^\ell \\ (c, \lambda) \leq c_0 \end{cases}, \quad (23)$$

где $I(\lambda)$ – произвольный критерий информативности.

Определение 5. Если $\lambda \in \Lambda^\ell$ и $(c, \lambda) \leq c_0$, то вектор λ называется планом задачи (23).

Через $\Lambda^\ell(c) = \{\lambda \in \Lambda^\ell: (c, \lambda) \leq c_0\}$ обозначим множество всех планов.

Пусть все компоненты $c = (c_1, c_2, \dots, c_N)$ вектора упорядочены, т.е. $c_i \leq c_{i+1}$ для любого i ($i = \overline{1, N-1}$). Используя компоненты этого вектора, построим следующие последовательности:

$$f_1 = \sum_{i=1}^{\ell} c_i, f_2 = \sum_{i=2}^{\ell+1} c_i, \dots, f_k = \sum_{i=k}^{\ell+k-1} c_i, \ell + k + 1 = N \Rightarrow k = N - \ell + 1.$$

Значит, при $z \geq 2$

$$f_z = f_1 + \sum_{i=1}^{z-1} (c_{\ell+i} - c_i).$$

Утверждение 6. $\Lambda^\ell(c) \neq \emptyset$ тогда и только тогда, когда $f_1 \leq c_0$.

Теперь найдем все k , удовлетворяющие условию $c_0 - f_k \geq 0$.

Примечание 3. Если $f_{N-\ell+1} \leq c_0$, то $\Lambda^\ell(c) \equiv \Lambda^\ell$.

Свойство 2. Если $c_0 - f_1 \geq 0$ и $c_0 - f_{N-\ell+1} < 0$, то существует такой p , $1 \leq p < N - \ell + 1$, для которого верно

$$\begin{cases} c_0 - f_p \geq 0 \\ c_0 - f_{p+1} < 0. \end{cases} \quad (24)$$

Утверждение 7. Если p является решением (24), то верно следующее $\Lambda_p^\ell = \{\lambda \in \Lambda^\ell: \lambda^j \in \{0,1\}, j = \overline{1, \ell + p - 1}, \lambda^i \equiv 0, i = \overline{\ell + p, N}\} \subseteq \Lambda^\ell(c)$.

Признак, имеющий наибольший индекс в векторе (c_1, c_2, \dots, c_N) и участвующий в ℓ информативных системах, определяется с помощью решения следующей системы неравенств:

$$\begin{cases} f_0 + c_j \leq c_0 \\ f_0 + c_{j+1} > c_0' \end{cases} \quad (25)$$

где $f_0 = \sum_{i=1}^{\ell-1} c_i, j = \overline{\ell + p, N}$.

Если c_q ($q < N$) является решением (25), то признаки, соответствующие компонентам $c_{q+1}, c_{q+2}, \dots, c_N$, исключаются из признакового пространства и осуществляется переход к q -мерному признаковому пространству.

Значит, $\Lambda^\ell(q) = \{\lambda: \sum_{j=1}^q \lambda^j = \ell\}$. Соответственно, $\Lambda^\ell(c)$ и Λ_p^ℓ примут следующий вид

$$\Lambda^\ell(c) = \{\lambda \in \Lambda^\ell(q): (c, \lambda) \leq c_0, c = (c_1, c_2, \dots, c_q)\},$$

$$\Lambda_p^\ell = \left\{ \lambda \in \Lambda^\ell(q): \sum_{j=1}^{\ell+p-1} \lambda^j = \ell, \sum_{j=\ell+p}^q \lambda^j = 0 \right\} \subseteq \Lambda^\ell(c).$$

В общем случае, каждому f_k соответствует единственный вектор $\lambda \in \Lambda^\ell(q)$, для которого верно следующее:

$$\lambda|_{f_k} = \left(\underbrace{0, 0, \dots, 0}_{k-1}, \underbrace{1, 1, \dots, 1}_\ell, \underbrace{0, 0, \dots, 0}_{q-(\ell+k-1)} \right).$$

Тогда для $f_{q-\ell+1}$ соответствует следующий вектор

$$\lambda|_{f_{q-\ell+1}} = \left(\underbrace{0, 0, \dots, 0}_{q-\ell}, \underbrace{1, 1, \dots, 1}_\ell \right).$$

Ниже приведен способ построения множества $\Lambda^\ell(c) \setminus \Lambda_p^\ell$ с помощью $\lambda|_{f_k}$.

Рекомендуется строить множества $\Lambda^\ell(c) \setminus \Lambda_p^\ell$ исходя из значения $q - p$, потому что, при построения этого множества количество операций будет меньше, если использовать $\lambda|_{f_{q-\ell+1}}$ при $q - p \leq 2\ell$, а $\lambda|_{f_k}$ ($k = \overline{1, p}$) при $q - p > 2\ell$.

Значит, построение множества $\Lambda^\ell(c) \setminus \Lambda_p^\ell$ осуществляется следующим образом:

При $q - p \leq 2\ell$ множество $\Lambda_{f_{q-\ell+1}}^\ell(c)$ относительно вектора $\lambda|_{f_{q-\ell+1}}$ определяется следующим образом:

$$\Lambda_{f_{q-\ell+1}}^\ell(c) = \{\lambda \in \Lambda^\ell(q): (c, \lambda) = (c, \lambda|_{f_{q-\ell+1}}) + \sum_{(i,j)} (c_i - c_j) \leq c_0,$$

$$i \in \{1, 2, \dots, \ell + p - 1\}, j \in \{\ell + p, \dots, q\}\} \subseteq \Lambda^\ell(c).$$

При $q - p > 2\ell$ множество $\Lambda_{f_k}^\ell(c)$ ($k = \overline{1, p}$) относительно вектора $\lambda|_{f_k}$ определяется следующим образом:

$$\Lambda_{f_k}^\ell(c) = \{\lambda \in \Lambda^\ell(q) \setminus \Lambda_p^\ell : (c, \lambda) = (c, \lambda|_{f_k}) + \sum_{(i,j)} (c_j - c_i) \leq c_0, \\ i \in \{1, 2, \dots, \ell + k - 1\}, j \in \{\ell + k, \dots, q\}\} \subseteq \Lambda^\ell(c).$$

Тогда получим задачи, эквивалентные задаче (23):

а) при $q - p \leq 2\ell$

$$\begin{cases} I_1(\lambda) \rightarrow \text{extr} \\ \lambda \in \Lambda_p^\ell \end{cases} \quad \text{и} \quad \begin{cases} I_{f_{q-\ell+1}}(\lambda) \rightarrow \text{extr} \\ \lambda \in \Lambda_{f_{q-\ell+1}}^\ell(c) \end{cases}$$

б) при $q - p > 2\ell$

$$\begin{cases} I_1(\lambda) \rightarrow \text{extr} \\ \lambda \in \Lambda_p^\ell \end{cases} \quad \text{и} \quad \begin{cases} I_{f_k}(\lambda) \rightarrow \text{extr} \\ \lambda \in \Lambda_{f_k}^\ell(c) \end{cases}$$

Понятно, что решение задачи (23) определяется следующим образом:
при $q - p \leq 2\ell$

$$I(\lambda) = \text{extr}\{I_1(\lambda), I_{f_{q-\ell+1}}(\lambda)\},$$

в противном случае

$$I(\lambda) = \text{extr}\left\{I_1(\lambda), \underset{f_k}{\text{extr}}\left(I_{f_k}(\lambda)\right)\right\}.$$

В параграфе 4.2 предложен метод и алгоритм ФОИС в условиях ограниченности ресурсов с использованием эвристического критерия информативности фишеровского типа. Если исключить признаки из признакового пространства, которые не могут участвовать при ФОИС, то задача (23) примет следующий вид:

$$\begin{cases} I(\lambda) = \frac{(a, \lambda)}{(b, \lambda)} \rightarrow \text{max} \\ \lambda \in \Lambda^\ell \\ (c, \lambda) \leq c_0 \end{cases}, \quad (26)$$

где $a, b, c \in R^q, a_j > 0, b_j > 0, c_j > 0, \lambda = (\lambda^1, \lambda^2, \dots, \lambda^q)$.

Введем предикат-функцию вида

$$s(\lambda) = \begin{cases} -1, & c_0 - (c, \lambda) < 0 \\ 1, & c_0 - (c, \lambda) \geq 0 \end{cases}, \quad \text{где } \lambda \in \Lambda^\ell.$$

Тогда (26) имеет следующий вид:

$$\begin{cases} I^*(\lambda) = \frac{(a, \lambda)}{(b, \lambda)} * s(\lambda) \rightarrow \text{max} \\ \lambda \in \Lambda^\ell \end{cases} \quad (27)$$

Из утверждения 1 следует $(\varphi(\lambda), \mu) \geq 0 \Leftrightarrow \frac{(a, \mu)}{(b, \mu)} \geq \frac{(a, \lambda)}{(b, \lambda)}$.

Отсюда вытекает следующее

Свойство 2. Если $(\varphi(\lambda), \mu) \geq 0$, то

- $I^*(\mu) \leq I^*(\lambda) < 0$ при $s(\lambda) = -1, s(\mu) = -1$;
- $I^*(\mu) \geq -I^*(\lambda) > 0 > I^*(\lambda)$ при $s(\lambda) = -1, s(\mu) = 1$;
- $I^*(\mu) \leq -I^*(\lambda) < 0 < I^*(\lambda)$ при $s(\lambda) = 1, s(\mu) = -1$;
- $I^*(\mu) \geq I^*(\lambda) > 0$ при $s(\lambda) = 1, s(\mu) = 1$.

Пусть $I^*(\mu) \geq I^*(\lambda)$, тогда

$$\frac{(a, \mu)}{(b, \mu)} s(\mu) \geq \frac{(a, \lambda)}{(b, \lambda)} s(\lambda) \Leftrightarrow (a, \mu)(b, \lambda)s(\mu) - (a, \lambda)(b, \mu)s(\lambda) \geq 0 \Leftrightarrow \\ \Leftrightarrow (a(b, \lambda)s(\mu) - b(a, \lambda)s(\lambda), \mu) \geq 0.$$

Из этого вытекает следующее

Свойство 3. Если $I^*(\mu) \geq I^*(\lambda)$, то

- a) $(\varphi(\lambda), \mu) \leq 0$ при $s(\lambda) = -1, s(\mu) = -1$;
- b) $(a(b, \lambda) + b(a, \lambda), \mu) \geq 0$ при $s(\lambda) = -1, s(\mu) = 1$;
- c) $(a(b, \lambda) + b(a, \lambda), \mu) \leq 0$ при $s(\lambda) = 1, s(\mu) = -1$;
- d) $(\varphi(\lambda), \mu) \geq 0$ при $s(\lambda) = 1, s(\mu) = 1$.

Из приведенных свойств видно, что если $s(\lambda) = -1, s(\mu) = 1$ для любого $\lambda, \mu \in \Lambda^\ell$ независимо от $(\varphi(\lambda), \mu)$, то всегда $I^*(\mu) > I^*(\lambda)$, если же $s(\lambda) = 1, s(\mu) = -1$, то всегда $I^*(\mu) < I^*(\lambda)$.

Примечание 4. Основная цель – найти оптимальные векторы для (27), и для этого необходимо определить, является ли следующий вектор оптимальным относительно заданного вектора λ . Для того, чтобы вектор μ был оптимальным, требуется выполнение условия $s(\mu) = 1$.

Примечание 5. Заданный вектор $\lambda \in \Lambda^\ell$ является планом задачи (27), если $s(\lambda) = 1$.

Проанализируем случай, когда $(\varphi(\lambda), \mu * s(\mu)) \geq 0$.

$$(\varphi(\lambda), \mu * s(\mu)) = (a(b, \lambda) - b(a, \lambda), \mu * s(\mu)) = (a, \mu)(b, \lambda)s(\mu) - \\ - (b, \mu)(a, \lambda)s(\mu) \geq 0 \Leftrightarrow \frac{(a, \mu)}{(b, \mu)} s(\mu) \geq \frac{(a, \lambda)}{(b, \lambda)} s(\mu) \Leftrightarrow I^*(\mu) \geq I(\lambda) * s(\mu)$$

Известно, что $\forall \lambda \in \Lambda^\ell$ верно $I(\lambda) \geq I^*(\lambda)$. Значит,

$$I^*(\mu) \geq I(\lambda) * s(\mu) \Leftrightarrow I^*(\mu) \geq I^*(\lambda) * s(\mu).$$

Отсюда вытекает следующее

Утверждение 8. Для $\forall \lambda, \mu \in \Lambda^\ell$ $I^*(\mu) \geq I^*(\lambda) * \frac{s(\mu)}{s(\lambda)}$ тогда и только тогда, когда $(\varphi(\lambda), \mu * s(\mu)) \geq 0$.

Пусть для заданного $\lambda \in \Lambda^\ell$ вектора $s(\lambda) = 1$.

Введем оператор

$$\mu: \Lambda^\ell \rightarrow \Lambda^\ell, \text{ где } (\varphi(\lambda), \mu * s(\mu)) = \max_{\substack{\eta \in \Lambda^\ell \\ s(\eta)=1}} (\varphi(\lambda), \eta(\lambda) * s(\eta))$$

Через $\mu(\lambda)$ обозначим μ удовлетворяющий данную условию.

Тогда для векторов λ и $\mu(\lambda)$ верно следующее

Свойство 4. $\forall \lambda \in \Lambda^\ell$ верно $(\varphi(\lambda), \mu(\lambda) * s(\mu)) \geq 0$, где $s(\lambda) = 1$.

Утверждение 9. Если $I^*(\lambda) = I^*(\mu(\lambda))$, то вектор $\mu(\lambda)$ является оптимальным решением для задачи (27).

В общем случае, количество оптимальных решений задачи (27) зависит от значений компонент вектора $\varphi(\lambda)$.

В параграфе 4.3 приведены некоторые свойства критерия информативности, т.е. функционала фишерского типа, использованной для

формирования ℓ информативных систем, а также для построения решающего правила.

В пятой главе диссертации «Исследование применения разработанных алгоритмов к модельным и прикладным задачам» изложены прикладное применение предложенных в диссертации методов и алгоритмов, а также правила формирования обучающей и контрольной выборок.

В параграфе 5.1 приведены описания IDEF0 и IDEF1 моделей программного комплекса «IBMT_S», разработанного на основе существующих и предложенных в диссертации методов и алгоритмов для решения задач интеллектуального анализа данных.

В параграфе 5.2 приведено правило формирования гипотетических обучающей и контрольной выборок для модельных задач интеллектуального анализа данных. Обоснована возможность формирования произвольного количества и объема обучающей и контрольной выборок с помощью данного правила.

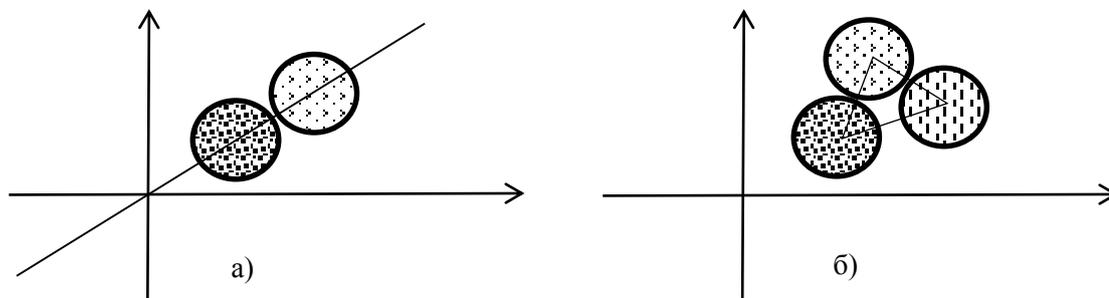


Рисунок 1. Гипотетический образ классов: а) количество классов - 2; б) количество классов – 3.

В параграфе 5.3 решены задачи выбора симптомокомплексов и диагностики для ишемической болезни сердца (ИБС). Здесь исследованы три вида ИБС: X_1 – «Прогрессирующая стенокардия», X_2 – «Острый-инфаркт миокарда» и X_3 – «Аритмическая форма». В каждом классе (виде) имеются, соответственно, по 140, 120 и 140 объектов (больных), а каждый объект описывается 62 разнотипными (количественными и качественными) признаками.

Для выбора симптомокомплексов применен алгоритм, предложенный в параграфе 3.2, а для компьютерной диагностики алгоритм, предложенный в параграфе 3.4 диссертации. В результате количество клинических признаков и симптомов для диагностики сократилось с 62 до 10. При $\ell = 10$ выявлено 7 симптомокомплексов, дающих экстремум значения заданному функционалу.

Таблица 1.

Наиболее информативные симптомокомплексы по ИБС

т/р	Выявленные симптомокомплексы
1.	$X_1, X_7, X_{16}, X_{22}, X_{27}, X_{36}, X_{41}, X_{51}, X_{57}, X_{62}$
2.	$X_1, X_7, X_{22}, X_{27}, X_{36}, X_{41}, X_{44}, X_{51}, X_{57}, X_{62}$
3.	$X_1, X_7, X_{22}, X_{27}, X_{36}, X_{41}, X_{47}, X_{51}, X_{57}, X_{62}$
4.	$X_1, X_7, X_{10}, X_{22}, X_{27}, X_{36}, X_{41}, X_{51}, X_{57}, X_{62}$
5.	$X_1, X_7, X_{22}, X_{27}, X_{34}, X_{36}, X_{41}, X_{51}, X_{57}, X_{62}$

6.	$X_1, X_2, X_7, X_{22}, X_{27}, X_{36}, X_{41}, X_{51}, X_{57}, X_{62}$
7.	$X_7, X_{22}, X_{27}, X_{34}, X_{36}, X_{41}, X_{47}, X_{51}, X_{57}, X_{62}$

В параграфе 5.4 приведены результаты применения программного комплекса IBMT_S для проведения экспериментального исследования изменения гидродинамических и гидрохимических показателей подземных вод (ПВ) на территории Ферганской области. Здесь решен ряд задач, в частности, исследованы изменения гидродинамических и гидрохимических показателей ПВ; определение основных источников загрязняющих компонентов, находящихся в подземных водах; выделения ореолы загрязнения подземных вод от различных источников загрязнения, находящихся на заданном участке и другие.

Ниже приведены результаты определения уровней загрязнения месторождения ПВ относительно минерализации. Здесь ПВ разделены на 6 классов по степени засоленности, т.е.

$$X = \bigcup_{p=1}^6 X_p, X_p \cap X_q = \emptyset, p \neq q, p, q = \overline{1,6},$$

где X_1 – незасоленные, $m_1 = 749$; X_2 – слабо засоленный, $m_2 = 294$; X_3 – средне засоленный, $m_3 = 178$; X_4 – соленый (солончак), $m_4 = 77$; X_5 – сильно засоленный, $m_5 = 151$; X_6 – очень сильно засоленный, $m_6 = 36$.

$x = (x^1, x^2, \dots, x^9) \in X$ – гидрогеологический объект, где x^1 – сухой остаток; x^2 – HCO_3 ; x^3 – CO_3 ; x^4 – NO_3 ; x^5 – Cl ; x^6 – SO_4 ; x^7 – Ca ; x^8 – Mg ; x^9 – $Na + K$.

Применяя программный комплекс «IBMT_S» для решения этой задачи, мы получим следующее.

Таблица 2.

Результаты ФОИС

ℓ	Без учета ресурса	$\max_{\lambda \in \Lambda^\ell} I(\lambda)$	С учетом ресурса	(с, λ)	$\max_{\lambda \in \Lambda^\ell} I(\lambda)$
4	1,2,3,6	4,3674	1,2,3,6	25	4,3674
5	1,2,3,4,6	4,1843	1,2,3,4,6	30	4,1843
6	1,2,3,4,5,6	4,0404	2,3,4,5,6,8	33	3,97
7	1,2,3,4,5,6,8	3,9513	Нет решения	$c > c_0$	

Ниже в таблице приведены результаты классификации по выбранному набору признаков.

Таблица 3.

Результаты классификации

ℓ	$\mathfrak{R}_{k-yq}(y, Y_p)$	$\mathfrak{R}_{BHA}(y, Y_p)$	$\mathfrak{R}_{XD}(y, Y_p)$	$\mathfrak{R}_{NT}(y, Y_p)$	$\mathfrak{R}_{Im}(y, Y_p)$
4	75,8	80,9	79,5	83,2	84,7
5	80	87,3	87,3	89,5	90
6	85,4	88,9	88	92	94

ЗАКЛЮЧЕНИЕ

Результаты проведенного диссертационного исследования на тему «Методы и алгоритмы формирования оптимальных информативных систем описания объектов в условиях ограниченности ресурсов» сводятся к следующим основным выводам.

1. Исследованы и проанализированы методы и алгоритмы, а также критерии информативности, используемые для решения задач интеллектуального анализа данных, в частности, задач классификации, кластеризации и формирования информативных систем описания объектов. Выявлено, что задачи классификации, кластеризации и снижения размерности признакового пространства в комплексном виде и в условиях ограниченности ресурсов мало изучены.

2. Формализованы задачи построения решающего правила и снижения размерности признакового пространства в задачах интеллектуального анализа данных. Это, в свою очередь, позволило разработать строго обоснованные методы и алгоритмы построения решающих правил и формирования информативных систем описания объектов с эффективным использованием развитого математического аппарата.

3. Обосновано, что метод ранжировки не всегда даёт оптимальные решения при формировании оптимальных информативных систем описания объектов. Выявлены достаточные условия оптимальности для метода ранжировки при формировании ℓ информативных систем с использованием критерия информативности фишеровского типа. Эти условия, в свою очередь, дали возможность определить условия применения данного метода и достоверность полученных результатов.

4. Разработаны усовершенствованный метод и алгоритм формирования оптимальных ℓ информативных систем описания объектов с использованием эвристических критериев информативности. Данный метод не зависит от типа признаков, описывающих объекты. Доказано, что полученные решения с помощью данного метода являются оптимальными, и процедура поиска оптимального решения заканчивается на 2-3 итерации. Обоснована эффективность данного метода в случае, когда оптимизационная задача имеет многочисленные критические точки. Это позволит эффективно использовать ресурсы при формировании оптимальных информативных систем.

5. Разработаны метод и алгоритм формирования оптимальных ℓ информативных систем, учитывающих коэффициент ошибок классификации. Это позволило повысить эффективность и надежность классификации, а также дало возможность выявлять и анализировать неправильно классифицированные объекты.

6. Разработан метод и алгоритм построения решающего правила, адекватного критерию информативности, использованного при формировании ℓ информативных систем описания объектов. С их помощью решены задачи классификации и снижения размерности признакового

пространства в комплексном виде. Они служат для повышения достоверности и эффективности классификации.

7. Разработаны метод и алгоритм формирования ℓ информативных систем с учетом ограниченности ресурсов, выделенных на измерение признаков объекта. В данном методе сначала строится такое множество ℓ информативных векторов λ , учитывающее ограниченность ресурсов на измерение признаков, что $\Lambda^\ell(c) = \{\lambda \in \Lambda^\ell: (c, \lambda) \leq c_0\}$. В этом подходе типы признаков и критерий информативности не являются существенными. Это даёт возможность решения оптимизационной задачи с помощью произвольного типа критерия информативности, который имеется в арсенале пользователя.

8. Для критерия информативности фишеровского типа предложены эффективный метод и алгоритм формирования оптимальных ℓ информативных систем при ограниченности ресурсов, выделенных на измерение признаков объекта. Для решения оптимизационной задачи формирования информативных систем признаков используются вектор-функция $\varphi(\lambda)$, указывающая направление наискорейшего роста функционала $I(\lambda)$, и предикат-функция $s(\lambda)$ указывающая на возможность участия набора признаков в формируемых информативных системах. Показана универсальность метода в отношении типа признаков. Решение данной оптимизационной задачи позволило информативно описать объекты наименьшим количеством признаков и с наименьшими затратами.

9. Выявлены свойства и особенности критерия информативности фишеровского типа, используемые в качестве конструктивной основы для формирования оптимальных информативных систем и построения решающего правила, а также осуществлена оценка ее эффективности. Показана возможность сокращения количества сравнений на $C_{N-\ell}^\ell$ при формировании оптимальных информативных систем с помощью метода полного перебора.

10. Усовершенствовано правило генерации модельной обучающей и контрольной выборок, на основе предъявляемых к ним требований в задачах интеллектуального анализа данных. Это дало возможность оценить надежность и эффективность разработанных методов и алгоритмов формирования информативных систем описания объектов и классификации в условиях ограниченности и неограниченности ресурсов, выделенных на измерение признаков объектов.

11. Разработан программный комплекс «IBMT_S» на базе существующих и предложенных в диссертации методов и алгоритмов. Программный комплекс внедрен для решения прикладных задач в сфере экономики, медицины, телекоммуникации, гидрогеологии и сельского хозяйства. Это дало возможность повысить производительность в среднем на 20-30% и снизить затраты в среднем на 10-15%.

**AD HOC SCIENTIFIC COUNCIL AT THE SCIENTIFIC COUNCIL
AWARDING SCIENTIFIC DEGREES DSc.13/30.12.2019.T.07.01 AT
TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES**

TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES

AKBARALIEV BAKHTIYORJON BAKIROVICH

**METHODS AND ALGORITHMS FOR THE FORMATION OF OPTIMAL
INFORMATIVE SYSTEMS FOR DESCRIBING OBJECTS IN
CONDITIONS OF LIMITED RESOURCES**

05.01.03 – Theoretical basis of computer science

**ABSTRACT
OF THE DOCTORAL (DSC) DISSERTATION OF TECHNICAL SCIENCES**

Tashkent – 2021

The theme of doctoral (DSc) dissertation of technical sciences was registered with the number of B2020.3.T370 at the Supreme Attestation Commission of the Cabinet of Ministers of the Republic of Uzbekistan.

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 Information and educational portal «ZiyoNet» website of www.ziynet.uz.

Scientific adviser:	Nishanov Akhram Khasanovich Doctor of Technical Sciences, Docent
Official opponents:	Muxamedieva Dilnoz Tulkunovna Doctor of Technical Sciences, Professor Gulyamov Shuhrat Mannapovich Doctor of Technical Sciences, Docent Nurmuxamedov Tolaniddin Ramziddinovich, Doctor of Technical Sciences, Docent
Leading organization:	National University of Uzbekistan named after Mirzo Ulugbek

The defense will take place « 17 » April 2021 at 14⁰⁰ on the meeting of Scientific council DSc.13/30.12.2019.T.07.01 at Tashkent University of Information Technologies (Address: 100084, Tashkent city, Amir Temur street, 108. Tel.: (99871) 238-64-43; fax: (99871) 238-65-52; e-mail: tuit@tuit.uz).

The dissertation is available at the Information Resource Centre of the Tashkent University of Information Technologies (is registered under № 199). (Address: 100084, Tashkent city, Amir Temur street, 108. Tel.: (99871) 238-65-44).

Abstract of dissertation sent out on « 17 » April 2021y.
(Dispatching protocol №. 1 on « 01 » April 2021 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

M.A.Rakhmatullaev

Chairman of academic seminar under the scientific council awarding scientific degrees, Doctor of Technical Sciences, Professor

INTRODUCTION (abstract of the dissertation of doctor of science (DSc))

The aim of research is to develop methods and algorithms for the formation of optimal informative systems, taking into account the allocated resource to determine the object based on heuristic criteria of informativeness.

The object of the research is the features (characteristics, description, properties, indicators) of objects (processes, phenomena, events) the values of which are presented in the form of tables "object-property" and "object-property-class".

The scientific novelty of the dissertation research is as follows:

determined sufficient conditions for the formation of the optimality of ℓ informative systems for the ranking method with respect to the criterion of informativeness of the Fisher type;

developed an improved method and algorithm for the formation of optimal ℓ informative systems using heuristic criteria of informativeness, taking into account the relationship between features, where objects are described by different types of features;

developed a method and an algorithm for the formation of optimal ℓ informative systems that increase the efficiency and reliability of classification, identify and analyze incorrectly classified objects based on the classification error rate;

developed a method and an algorithm for constructing a decision rule in the classification problem based on the corresponding criterion of information content of the formation of optimal ℓ informative systems;

developed a method and an algorithm for the formation of informative systems for describing objects, taking into account resources with respect to the criterion of informativeness of an arbitrary type;

developed an effective method and algorithm for the formation of optimal ℓ informative systems with limited resources allocated for determining the object based on the heuristic criterion of information content of the Fisher type;

improved the rule for generating model training and control samples based on the requirements for samples in data mining tasks.

Implementation of the research results. Based on the obtained scientific results on the methods and algorithms for the formation of optimal informative systems for describing objects in conditions of limited resources:

the method and algorithm for the formation of optimal ℓ informative systems, which increases the efficiency and reliability of classification, identifies and analyzes incorrectly classified objects based on the classification error rate, the method and algorithm for the formation of optimal ℓ informative systems in conditions of limited resources for determining the object, based on the Fisher's heuristic criterion of informativeness type, as well as the method and algorithm for constructing the decision rule, and the software complex "IBMT_S" developed on their basis, have been introduced into the activities of the State Enterprise Institute of Hydrogeology and Engineering Geology (certificate of the State Committee of Geology and Mineral Resources of the Republic of Uzbekistan No. 03/07 dated

August 25, 2020 No.). The results of scientific research allowed to increase labor productivity by an average of 20-30%, to reduce various costs by 10-15% due to the classification and clustering of hydrogeological objects;

an improved method and algorithm for the formation of optimal informative systems using heuristic criteria of informativeness, taking into account the relationship between features, where objects are described by different types of features, and the software complex "IBMT_S" developed on their basis has been introduced into the activities of the State Unitary Enterprise Republican Center of Telecommunication Networks of Uzbekistan (certificate of the Ministry for Development of Information Technologies and Communication of the Republic of Uzbekistan №33-8/5595 of September 25, 2020). The results of the scientific research made it possible to increase labor efficiency by an average of 20-30%, to reduce the time and other resources by 40% by defining informative parameters describing the technical condition of telecommunication networks, as well as classifying telecommunication networks and users;

the method and algorithm for the formation of informative systems for describing objects, taking into account resources regarding the criterion of informativeness of an arbitrary type, and the software complex "IBMT_S" developed on their basis are introduced into the activities of the Agency for the Management of State Assets of the Republic of Uzbekistan (certificate of the Ministry for Development of Information Technologies and Communication of the Republic of Uzbekistan №33-8/5595 of September 25, 2020). The results of scientific research allowed to increase labor productivity by an average of 10-15%, reduce workload and financial costs by 10-20% and 5-8%, respectively.

Structure and scope of the dissertation. The dissertation consists of the introduction, five chapters, conclusion, the list of used literature and appendix. The dissertation volume is 184 pages.

ЭЪЛОН ҚИЛИНГАН ИШЛАР РЎЙХАТИ
СПИСОК ОПУБЛИКОВАННЫХ РАБОТ
LIST OF PUBLISHED WORKS

I бўлим (I часть; I part)

1. М.М.Камилов, А.Х.Нишанов, Б.Б.Акбаралиев. Об одном методе построения решающего правила при решении основных задач распознавания образов//Кимёвий технология, назорат ва бошқарув, Илмий-техникавий журнал, 3-сон, 2006, 59-67 б. (05.00.00; №12)

2. Б.Б.Акбаралиев. Формирование информативных наборов признаков в сложных системах распознавания//ТАТУ хабарлари, 2-сон, 2007, 47-50 б. (05.00.00; №31)

3. Т.С.Гаипназаров, М.М.Камилов, Б.Б.Акбаралиев. Интеллектуализация систем поддержки принятия решений//ТАТУ хабарлари, 3-сон, 2008, 9-14 б. (05.00.00; №31)

4. Б.Б.Акбаралиев, О.Ж.Бобомурадов. Об одном подходе обработки данных в задачах распознавания образов//«Информатика ва энергетика муаммолари» Ўзбекистон журналы, 5-сон, 2008, 20-23 б. (05.00.00; №5)

5. Т.Гаипназаров, М.М.Камилов, Б.Б.Акбаралиев. Разработка информационной системы поддержки принятия решений для управления учебным процессом//ТАТУ хабарлари, 4-сон, 2009, 97-102 б. (05.00.00; №31)

6. Б.Б.Акбаралиев, А.Т.Рахмонов, А.К.Эргашев. Об одном методе сокращения размерности объема выборки в интеллектуальном анализе данных//Информатика ва Энергетика муаммолари Ўзбекистон журналы, 1-сон, 2011, 76-79 б. (05.00.00; №5)

7. А.Т.Рахманов, Б.Б.Акбаралиев, О.Б.Рузибаев, У.А.Хасанов. Об одном модифицированном способе решения задачи классификации//Кимёвий технология, назорат ва бошқарув, Халқаро илмий-техникавий журнал, 4-сон, 2014, 85-91 б. (05.00.00; №12)

8. А.Х.Нишанов, Б.Б.Акбаралиев, О.Б.Рузибаев, О.К.Хужаев. Сравнительный анализ алгоритмов на основе нечеткого К-средних с применением различных метрик //Кимёвий технология, назорат ва бошқарув, Халқаро илмий-техникавий журнал, 6-сон, 2014, 78-82 б. (05.00.00; №12)

9. Б.Б.Акбаралиев, О.К.Хужаев, О.Б.Рузибаев, Ш.Н.Саидрасулов. Moodle тизими базаси элементлари кесимида Data Mining усулларида фойдаланилган ҳолда ўқитувчи фаолиятини баҳолаш//ТАТУ хабарлари, 2(34)-сон, 2015, 11-16 б. (05.00.00; №31)

10. А.Х.Nishanov, E.Sh. Avazov, B.B.Akbaraliyev. Partial Selection Method and Algorithm for Determining Graph-Based Traffic Routes in a Real-Time Environment//International Journal of Innovative Technology and Exploring Engineering (IJITEE), Vol.-8 Issue-6S, 2019, pp.696-698 (№ 3, Scopus, IF=0.6)

11. А.Х.Нишанов, Б.Б.Акбаралиев, Ф.П.Жўраев. Юрак ишемик касалликлари симптокомплексларини танлашнинг мослашувчан тасодифий

кидирувга асосланган алгоритми//Муҳаммад ал-Хоразмий авлодлари, Илмий-амалий ва ахборот таҳлилий журнал, 2(12)-сон, 2020, 10-15 б. (05.00.00; №10)

12. Б.Б.Акбаралиев, А.Х.Нишанов. Об эффективном способе оптимизации метода ранжировки для критериев информативности фишеровского типа//ТАТУ хабарлари, 4-сон, 2020, 11-16 б. (05.00.00; №31)

13. В.В.Акбаралиев. An effective method for attribute subset selection, considering the resource in pattern recognition// Chemical Technology. Control and Management. International scientific and technical journal. Iss. 4, (94), 2020, pp.57-64. (05.00.00; №12)

14. А.Кх.Нисанов, В.В.Акбаралиев, Г.П.Жураев, М.А.Кхасанова, М.Кх.Максудова, З.Ф.Умарова. The Algorithm for Selection of Symptom Complex of Ischemic Heart Diseases Based on Flexible Search//Journal of Cardiovascular Disease Research, Vol 11, Issue 2, 2020, p.218-223 (№ 3, Scopus, IF=0.4).

15. А.Кх.Нисанов, В.В.Акбаралиев, В.С.Самандаров, О.К.Ахмедов, Ш.Кх.Тажобаев. An algorithm for classification, localization and selection of informative features in the space of polytypic data//Webology, Volume 17, № 1, 2020, p.341-364. (№ 3, Scopus, IF=0.9)

II бўлим (II часть; II part)

16. М.М.Камилов, А.Н.Нисанов, В.В.Акбаралиев. Methods of forming of optimal sign space for object recognition in the class of logic-heuristic algorithms//Fourth World Conference on Intelligent Systems for Industrial Automation - WCIS 2006, Tashkent, pp. 45-52.

17. М.М.Камилов, А.Н.Нисанов, В.В.Акбаралиев. About one clustering algorithm in intellectual data analysis// Proceedings of ICEIC 2008, Tashkent, 2008, pp. 476-478.

18. Т.Ф.Бекмуратов, О.Ж.Бобомуродов, В.В.Акбаралиев, Н.Рахимов. Organizing of teaching materials in the intelligent teaching systems//Proceedings of ICEIC 2008, Tashkent, 2008, pp. 369-371.

19. М.М.Камилов, Б.Б.Акбаралиев, О.Ж.Бобомуратов. Эвристический метод построения информативного признакового пространства в интеллектуальных системах анализа данных алгоритмами распознавания// Интеллектуальные системы: Труды Восьмого международного симпозиума/под ред. К.А.Пупкова, –М.: РУСАКИ, 2008, – С.113-116.

20. М.М.Камилов, В.В.Акбаралиев, О.Ж.Бобомуродов, Р.Н.Кхолиқназаров. λ -self-descriptiveness and formation of irredandant description of objects recognition//Fifth World Conference on Intelligent Systems for Industrial Automation - WCIS 2008, Tashkent, 2008, pp.169-171.

21. Б.Б.Акбаралиев, Н.А.Эгамбердиев, Ш.Э.Эргашев. Икки ўлчовли динамик объектларни нейробошқариш//«Инновацион ғоялар, технологиялар ва лойиҳаларни ишлаб чиқаришга тадбиқ этиш муаммолари» IV Республика илмий-техник конференцияси илмий ишлар тўплами, 2012, 35-37 б.

22. Б.Б.Акбаралиев, А.О.Ашуров. Тимсолларни таниб олиш масалаларида белгилар мажмуасининг муҳимлилик даражаларини аниқлаш усуллари//«Инновацион ғоялар, технологиялар ва лойиҳаларни ишлаб чиқаришга тадбиқ этиш муаммолари» IV Республика илмий-техник конференцияси илмий ишлар тўплами, 2012, 37-39 б.

23. Б.Б.Акбаралиев, З.Юсупова, А.Ғ.Маджидов. Ҳал қилувчи қоидани қуриш алгоритми ҳақида//«Ахборот технологиялари ва телекоммуникация муаммолари» ёш олимлар, тадқиқотчилар, магистрант ва талабаларнинг Республика илмий-техник конференцияси маърузалар тўплами, I том, Тошкент, 2012, 22 б.

24. Б.Б.Акбаралиев, У.И.Рахмонов, У.А.Хасанов. Тасвирларни аниқлашда белгилар фазосини шакллантириш эҳтиёж чеклангандаги усули//«XXI аср – интеллектуал авлод асри» шиори остида ёш олимлар ва талабаларнинг худудий илмий-амалий конференция, Нукус, 2014, 267-268 б.

25. О.А.Мамарауфов, Б.Б.Акбаралиев. Требования к геометрии камеры для распознавания динамического объектов на видеоизображениях// Материалы международной научно-технической конференции «Радиотехника, телекоммуникация и информационные технологий: Развития перспективы и проблемы», Ташкент – 2015. – С. 174-177.

26. M.M.Kamilov, Z.B.Mingliqulov, V.B.Akbaraliev. Using heuristic algorithms for solving the task of combinatorial optimization//EIGHTH World Conference on Intelligent Systems for Industrial Automation - WCIS 2014, Tashkent, pp.9-13.

27. Б.Б.Акбаралиев Катта хажмдаги маълумотларни қайта ишлаш йўллари//«Иқтисодиётнинг реал тармоқларини инновацион ривожланишида ахборот-коммуникация технологияларининг аҳамияти» Республика илмий-техника анжумани, Тошкент, 2020, 140-144 б.

28. A.Kh.Nishanov, V.B.Akbaraliev, Sh.Kh.Tajibaev. About one feature selection algorithm method in pattern recognition//Eleventh World Conference on Intelligent Systems for Industrial Automation - WCIS 2020, Tashkent.

29. A.Kh.Nishanov, V.B.Akbaraliev, G.P.Juraev. A symptom selection algorithm based on classification errors//International Conference on Information Science and Communications Technologies ICISCT 2020 Applications, Trends and Opportunities, Tashkent, 2020.

30. А.Х.Нишанов, Б.Б.Акбаралиев, А.Х.Холиқназаров, О.Ахмедов, Ш.Тажобаев. Ер ости сув ҳолатларини аниқлашда муҳим белгиларни танлаш алгоритми//International scientific online conference «Actual problems and prospectes of the development of intelligent information and communication systems», PCS-2020, Tashkent, 2020, 181-185-б.

31. А.Х.Нишанов, Б.Б.Акбаралиев, А.Х.Холиқназаров, О.Ахмедов, Ш.Тажобаев. Ер ости сувлари давлат мониторинг объектларини интеллектуал таҳлил қилиш// International scientific online conference «Actual problems and prospectes of the development of intelligent information and communication systems», PCS-2020, Tashkent, 2020, 185-190-б.

32. А.Х.Нишанов, Б.Б.Акбаралиев, О.Б.Рузибоев, Б.С.Самандаров, И.В.Хан. Программное обеспечение для анализа качества электронных информационных ресурсов и мониторинга работы системы виртуального образования Moodle// Свидетельство о депонировании объектов интеллектуальной собственности. Avtor.uz регистрационной №2173. Ташкент, 2014.

33. О.А.Мамарауфов, Б.Б.Акбаралиев, Т.А.Кучкаров. Камера курсор// Свидетельство об официальном регистрации программы для электронно-вычислительных машин. №DGU03430, Ташкент, 2015.

34. А.Х.Нишанов, Б.Б.Акбаралиев, Ф.П.Жўраев, О.Б.Рўзибоев, Ж.Зоиров. Тиббий ташхис ҳолатларини баҳолашнинг ҳисоблаш алгоритмларига асосланган дастури// Свидетельство об официальном регистрации программы для электронно-вычислительных машин. №DGU08573, Ташкент, 2020.

35. Б.Б.Акбаралиев, О.К.Хўжаев, Р.Х.Холиқназаров. Объектларни тавсифловчи информатив тизимларни шакллантирувчи ва синфлаштирувчи дастурий восита («IBMT_S»)// Ўзбекистон Республикаси Адлия вазирлиги ҳузуридаги Интеллектуал мулк агентлиги. ЭҲМ учун яратилган дастурнинг расмий рўйхатдан ўтказилганлиги тўғрисидаги гувоҳнома. № DGU 10033, Тошкент, 2021.

Автореферат «Муҳаммад ал-Хоразмий авлодлари» Илмий-амалий ва ахборот таҳлилий журнали таҳририятида таҳрирдан ўтказилди ҳамда ўзбек, рус ва инглиз тилларидаги матнлар ўзаро мувофиқлаштирилди.