

**ТОШКЕНТ АХБОРОТ ТЕХНОЛОГИЯЛАРИ УНИВЕРСИТЕТИ
ҲУЗУРИДАГИ ИЛМий ДАРАЖАЛАР БЕРУВЧИ
DSc.27.06.2017.Т.07.01 РАҚАМЛИ ИЛМий КЕНГАШ
АСОСИДАГИ БИР МАРТАЛИК ИЛМий КЕНГАШ**

**МУҲАММАД АЛ-ХОРАЗМий НОМИДАГИ ТОШКЕНТ АХБОРОТ
ТЕХНОЛОГИЯЛАРИ УНИВЕРСИТЕТИ ҲУЗУРИДАГИ
АХБОРОТ-КОММУНИКАЦИЯ ТЕХНОЛОГИЯЛАРИ
ИЛМий-ИННОВАЦИОН МАРКАЗИ**

МАМАТОВ НАРЗИЛО СОЛИДЖОНОВИЧ

**МАЪЛУМОТЛАРГА ДАСТЛАБКИ ИШЛОВ БЕРИШ УСУЛ ВА
АЛГОРИТМЛАРИ**

05.01.03 – Информатиканинг назарий асослари

**ТЕХНИКА ФАНЛАРИ ДОКТОРИ (DSc)
ДИССЕРТАЦИЯСИ АВТОРЕФЕРАТИ**

Тошкент – 2018

Докторлик (DSc) диссертацияси автореферати мундарижаси

Оглавление автореферата докторской (DSc) диссертации

Contents of the Doctoral (DSc) Dissertation Abstract

Маматов Нарзилло Солидждонович

Маълумотларга дастлабки ишлов бериш усул ва алгоритмлари.....3

Маматов Нарзилло Солидждонович

Методы и алгоритмы предварительной обработки данных.....29

Mamatov Narzillo Solidjonovich

Methods and algorithms for preliminary data processing55

Эълон қилинган ишлар рўйхати

Список опубликованных работ

List of published works.....59

**ТОШКЕНТ АХБОРОТ ТЕХНОЛОГИЯЛАРИ УНИВЕРСИТЕТИ
ҲУЗУРИДАГИ ИЛМИЙ ДАРАЖАЛАР БЕРУВЧИ
DSc.27.06.2017.Т.07.01 РАҚАМЛИ ИЛМИЙ КЕНГАШ
АСОСИДАГИ БИР МАРТАЛИК ИЛМИЙ КЕНГАШ**

**МУҲАММАД АЛ-ХОРАЗМИЙ НОМИДАГИ ТОШКЕНТ АХБОРОТ
ТЕХНОЛОГИЯЛАРИ УНИВЕРСИТЕТИ ҲУЗУРИДАГИ
АХБОРОТ-КОММУНИКАЦИЯ ТЕХНОЛОГИЯЛАРИ
ИЛМИЙ-ИННОВАЦИОН МАРКАЗИ**

МАМАТОВ НАРЗИЛЛО СОЛИДЖОНОВИЧ

**МАЪЛУМОТЛАРГА ДАСТЛАБКИ ИШЛОВ БЕРИШ УСУЛ ВА
АЛГОРИТМЛАРИ**

05.01.03 – Информатиканинг назарий асослари

**ТЕХНИКА ФАНЛАРИ ДОКТОРИ (DSc)
ДИССЕРТАЦИЯСИ АВТОРЕФЕРАТИ**

Тошкент – 2018

Техника фанлари бўйича фан доктори (DSc) диссертацияси мавзуси Ўзбекистон Республикаси Вазирлар Маҳкамаси ҳузуридаги Олий аттестация комиссиясида В2017.2.DSc/T86 рақам билан рўйхатга олинган.

Диссертация Тошкент ахборот технологиялари университети ҳузуридаги ахборот-коммуникация технологиялари илмий-инновацион марказида бажарилган.

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

Илмий маслаҳатчи:

Фозилов Шавкат Хайруллаевич
техника фанлари доктори, профессор

Расмий оппонентлар:

Кобулов Анвар Василович
техника фанлари доктори, профессор

Рустамов Насим Тулагенович
техника фанлари доктори, профессор (Қозоғистон)

Нурмухамедов Толаниддин Рамзиддинович
техника фанлари доктори

Етакчи ташкилот:

Тошкент давлат техника университети

Диссертация ҳимояси Тошкент ахборот технологиялари университети ҳузуридаги DSc.27.06.2017.T.07.01 рақамли Илмий кенгашнинг 2018 йил «___» _____ соат ___ даги мажлисида бўлиб ўтади. (Манзил: 100202, Тошкент шаҳри, Амир Темур кўчаси, 108-уй. Тел.: (99871) 238-64-43, факс: (99871) 238-65-52, e-mail: tuit@tuit.uz).

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

Диссертация автореферати 2018 йил «___» _____ куни тарқатилди.
(2018 йил «___» _____ даги ___ рақамли реестр баённомаси.)

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

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

М.М.Қамилов
Илмий даражалар берувчи илмий кенгаш
кошидаги илмий семинар раиси, т.ф.д., академик

КИРИШ (фан доктори (DSc) диссертацияси аннотацияси)

Диссертация мавзусининг долзарблиги ва зарурати. Жаҳонда ахборотнинг компьютерли таҳлили, маълумотларни таснифий қайта ишлаш ва тимсолларни таниб олиш усуллари ишлаб чиқишга катта эътибор қаратилмоқда. Ушбу йўналишларда таниб олиш ва тавсифлаш тизимлари, техник ташхислаш, робототехник тизимлар, материалшунослик, технологик жараёнлар, саноат объектларини бошқариш каби техник ва саноат иловаларини ишлаб чиқиш муҳим масалаларидан бири бўлиб қолмоқда. Ҳозирги кунда таниб олиш ва таснифлашнинг детерминлашган, эҳтимолий, мантиқий, тузилмавий каби кўплаб усуллари ишлаб чиқилган. Хорижий мамлакатларда, жумладан АҚШ, Германия, Россия, Хитой, Япония, Жанубий Корея ва бошқа мамлакатларда тимсолларни таниб олиш ва таснифлаш тизимларини назарий ва амалий масалаларини ечишга катта эътибор қаратилмоқда.

Жаҳонда объект, ҳодиса ва жараёнларни бошқариш, башоратлаш, идентификациялаш масалаларини ечишга йўналтирилган ва таниб олиш алгоритмларига асосланган сунъий интеллект тизимларини яратишга қаратилган кенг қамровли илмий тадқиқотлар олиб борилмоқда. Бу борада, жумладан, маълумотлардаги бўшлиқларни тўлдириш, ҳалақит ва сакрашларни бартараф этиш, ўзгарувчилар типларини унификациялаш усул ва алгоритмларини такомиллаштириш, маълумотларга дастлабки ишлов бериш ва қайта ишлаш усуллари ишлаб чиқиш, объектларни информатив таснифлаш, таниб олиш, информатив белгиларни танлаш, ноинформатив белгиларни аниқлаш усул ва алгоритмларини яратиш муҳим вазифалардан ҳисобланади.

Республикамизда иқтисодий, ижтимоий, халқ хўжалиги ва бошқа соҳаларда корпоратив бошқаришга ахборот-коммуникация технологиялари қўлланилмоқда. Ушбу йўналишда маълумотларга дастлабки ишлов бериш ва компьютерли таҳлил қилишга, жумладан, объект, ҳодиса ва жараёнларнинг информатив таснифини шакллантириш усул ва алгоритмларини ишлаб чиқишга алоҳида эътибор қаратилмоқда. 2017-2021 йилларда Ўзбекистон Республикасини янада ривожлантириш бўйича Ҳаракатлар стратегиясида, жумладан «... корпоратив бошқарувнинг замонавий стандарт ва усуллари жорий этиш, ... иқтисодиёт, ижтимоий соҳа, бошқарув тизимида ахборот-коммуникация технологияларини жорий этиш»¹ вазифалари белгиланган. Мазкур вазифаларни амалга ошириш, жумладан, тимсолларни таниб олиш усулларида фойдаланиб қарор қабул қилишни қўллаб-қувватлаш тизимларини яратиш муҳим масалалардан бири ҳисобланади. Бу борада, қарор қабул қилишда фойдаланиладиган бошланғич маълумотларга дастлабки ишлов бериш усуллари такомиллаштириш муҳим аҳамият касб этади.

¹Ўзбекистон Республикаси Президентининг 2017 йил 7 февралдаги ПФ-4947-сон “Ўзбекистон Республикасини янада ривожлантириш бўйича Ҳаракатлар стратегияси тўғрисида”ги Фармони

Ўзбекистон Республикаси Президентининг 2017 йил 7 февралдаги ПФ-4947-сон “Ўзбекистон Республикасини янада ривожлантириш бўйича Харажатлар стратегияси тўғрисида”ги, 2017 йил 29 ноябрдаги ПФ-5264-сон “Ўзбекистон Республикаси инновацион ривожлантириш вазирлигини ташкил этиш тўғрисидаги” Фармонлари, 2012 йил 21 мартдаги ПҚ-1730-сон «Замонавий ахборот-коммуникация технологияларини янада жорий этиш ва ривожлантириш чора-тадбирлари тўғрисида»ги, 2014 йил 3 апрелдаги ПҚ-2158-сон «Ахборот-коммуникация технологияларини иктисодиётнинг реал секторига янада жорий этиш чора-тадбирлари тўғрисида»ги, 2017 йил 18 апрелдаги ПҚ-2898-сон «Ички ишлар органларининг жиноятларни тергов қилиш соҳасидаги фаолиятини тубдан такомиллаштириш чора-тадбирлари тўғрисида»ги Қарорлари ҳамда мазкур фаолиятга тегишли бошқа меъёрий-ҳуқуқий ҳужжатларда белгиланган вазифаларни амалга оширишга ушбу диссертация тадқиқоти муайян даражада хизмат қилади.

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

Диссертация мавзуси бўйича хорижий илмий-тадқиқотлар шарҳи². Тимсолларни таниб олиш тизимларини яратиш ва усулларини ишлаб чиқиш ҳамда тадқиқ қилинаётган объектлар информатив тавсифини шакллантиришга қаратилган кенг қамровли илмий тадқиқотлар жаҳоннинг етакчи илмий марказлари ва олий таълим муассасалари, жумладан Carnegie Mellon University, University of California, Department of Electrical and Computer Engineering University of Texas (АҚШ), Darmstadt University of Applied Sciences (Германия), Attar Software Ltd (Буюк Британия), SIPINA, University of Lyon (Франция), Izmir University (Туркия), Institute of Cybernetics of National Academy of Sciences of Ukraine (Украина), United Institute of Informatics Problems, Belarusian State University of Informatics and Radioelectronics (Белоруссия), Azerbaijan Association of Zadeh’s Legacy, Department of EECS, Department of Business Administration (Озарбайжон), Россия Фанлар Академиясининг Ҳисоблаш маркази, Ахборотларни узатиш муаммолари институти, Санкт-Петербург информатика ва автоматлаштириш институти, Москва Давлат Университети, Амалий математика институти, Томск Давлат Университети (Россия Федерацияси), Автоматлаштириш институти (Хитой), Tokyo Institute of Technology (Япония), Ўзбекистон Миллий университети (Ўзбекистон) кенг қамровли илмий тадқиқот ишлари олиб борилмоқда.

Маълумотларга дастлабки ишлов бериш, тимсолларни аниқлаш

²Диссертация мавзуси бўйича илмий тадқиқотлар шарҳи <https://elibrary.ru/item.asp?id=21681096>, <https://studfiles.net/1376730/>, <https://cyberleninka.ru/.../obzor-metodov-raspoznavaniya-simvolov>, <http://intellect-tver.ru/?p=165>, <http://www.nickart.spb.ru/analysis/market.php>, <http://www.cedar.buffalo.edu/~srihari/CSE555/>, <https://www.ibm.com/developerworks/ru/library/ba-data-mining-techniques/index.html>, www.open.edu/openlearn-create/mod/resource/view.php?id, <http://ieeexplore.ieee.org/document/824819/> ва бошқа манбалар асосида ишлаб чиқилган.

усуллари ва алгоритмларини яратиш ҳамда таниб олиш ва тавсифлаш тизимларини такомиллаштиришга оид жаҳонда олиб борилган тадқиқотлар натижасида қатор, жумладан қуйидаги илмий натижалар олинган: маълумотлардаги халақитларни бартараф этиш усул ва алгоритмлари яратилган (Международная академия информатизации, Россия Федерацияси; Stanford University, АҚШ; Transport and telecommunication institute, Латвия); маълумотларни силлиқлаш ва филтрлаш усуллари ва алгоритмлари ишлаб чиқилган (Massachusetts Institute of Technology, International Institute of Forecasters, АҚШ; Центральный экономико-математический институт, Россия Федерацияси); объектни характерловчи белгиларни бир хил типга ўтказиш усул ва алгоритмлари ишлаб чиқилган (Московский государственный технический университет имени Н.Э.Баумана, Санкт-Петербургский институт информатики и автоматизации, Россия Федерацияси); белгилар фазосининг ўлчамини камайтириш усуллари ва алгоритмлари ишлаб чиқилган (Институт проблем управления, Россия Федерацияси; Bell Labs, АҚШ; NEC Labs, Япония); нутқни таниб олиш учун дастурий таъминот яратилган (Nuance Communications, Google, АҚШ; Yandex, Россия Федерацияси); шахсни таниб олиш тизимлари яратилган (Kairos, Англия; NVIDIA, Microsoft, АҚШ); матнни таниб олиш тизимлари яратилган (АВВУУ, Россия Федерацияси).

Дунёда маълумотларни таҳлил қилиш ва тимсолларни таниб олиш усул ва алгоритмларини такомиллаштириш ва янгиларини яратиш бўйича қатор, жумладан қуйидаги устувор йўналишларда тадқиқотлар олиб борилмоқда: дастлабки ахборотнинг ўта чекланган ҳажмида ишловчи ва таниб олиш тизимлари интеллектуаллигини таъминловчи усул ва алгоритмлар ишлаб чиқиш; катта ўлчамли белгилар фазосида таниб олиш усул ва алгоритмларини ишлаб чиқиш; тасвирлардаги объектларни аниқлаш усул ва алгоритмлар ишлаб чиқиш; кирувчи ахборотнинг интеграциялашган қайта ишлашини таъминлаб, “юмшоқ” ҳисоблашлар технологияларидан фойдаланган ҳолда нутқ, матн, шахсни идентификация қилиш каби тизимларини яратиш.

Муаммонинг ўрганилганлик даражаси. Маълумотларни таҳлил қилиш, хусусан тимсолларни таниб олиш масалаларида белгилар фазоси ўлчамини камайтиришнинг илмий асослари С.А.Айвазян, Б.Н.Адасовский, В.В.Александров, С.В.Абламейко, В.И.Васильев, Н.Д.Горский, Н.Г.Загоруйко, Г.С.Лбов, Н.А.Игнатъев, Ш.Ю.Радис, В.В.Старовойтов, I.Marill, D.M.Green ва бошқа олимлар томонидан яратилган ва уларнинг ишларида баён этилган. Ю.И.Журавлев, М.М.Камилов, З.Т.Адылова, Р.А.Лутфуллаев, Ш.Е.Тулягановларнинг ишларида дастлабки белгининг муҳимлик ўлчови хоссалари тадқиқ этилган, бунда белгининг муҳимлик ўлчови белгини олиб ташлашда тегишли равишда қайта ишланган “овозлар” деб аталувчиларни камайтириш даражасини ифодалайди. Ушбу ўлчов усул муаллифлари томонидан белгининг ахборот оғирлиги деб аталган.

А.Л.Горелик, Г.И.Кутин, К.А.Чепонис, Д.А.Жвинерайте, Б.С.Бусыгин, Л.В.Мирошниченко, R.Fisher каби олимларнинг ишлари белгилар

информативлигининг эвристик мезонлари тадқиқи ҳамда бу мезонлар асосида объектларнинг информатив тавсифини шакллантиришга мўлжалланган танлов усулларини ишлаб чиқишга бағишланган. Кейинчалик бу ёндошув М.М.Камилов, Ш.Х.Фозилов, А.Х.Нишанов томонидан ривожлантирилган бўлиб, уларнинг ишларида информатив белгилар мажмуасини аниқлашнинг қидирув усуллари таклиф этилган ва белгиларнинг ℓ -информатив вектори тушунчаси киритилган. Р.Ҳамдамовнинг ишларида Фишер эвристик информативлик мезони асосида информатив белгиларни аниқлашнинг аналитик усуллари биринчилардан бўлиб таклиф этилган.

Тимсолларни таниб олишда белгилар фазоси ўлчамини камайтириш масаласини ечиш билан боғлиқ тадқиқотларнинг деярли барчаси ҳал этиладиган масала учун энг мақбул мезон ва усул жуфтлигини танлашга келтирилади, яъни ҳисоблаш сарф-харажатлари кам ва таниб олиш сифати юқори бўлган мезон ва усул жуфтлигини танлаш талаб этилади. Бунда белгилар информативлигининг кенг тарқалган мезонларига мос математик асосланган усулларни яратиш муаммоси долзарб бўлиб, ҳозирги кунда етарли даражада ўрганилмаган.

Диссертация тадқиқотининг диссертация бажарилган илмий-тадқиқот муассасасининг илмий-тадқиқот ишлари режалари билан боғлиқлиги. Диссертация тадқиқотлари Ахборот-коммуникация технологиялари илмий-инновацион маркази илмий тадқиқот режасининг ФА-А17-Ф010: «Шахсни биометрик идентификация қилишнинг кўп поғонали тизимининг алгоритмик ва дастурий таъминотини ишлаб чиқиш» (2009-2011); А5-ФА-Ф022: «Биометрик ечимларнинг гибрид алгоритмлари ва дастурий таъминотини ишлаб чиқиш ва амалиётда қўллаш» (2012-2014); А5-003: «Кўллёзма матнларини таҳлил қилиш тизимининг алгоритмик таъминотини яратиш ва амалда қўллаш» (2015-2017); БВ-М-Ф4-003: «Белгиларнинг ўзаро боғлиқликларини баҳолашга асосланган таниб олиш операторларини синтез қилиш усуллари» (2017-2020) мавзулардаги лойиҳалари доирасида бажарилган.

Тадқиқотнинг мақсади эвристик информативлик мезонларининг хусусий ва умумий кўринишларидан фойдаланган ҳолда объектлар информатив тавсифини шакллантиришнинг усул ва алгоритмларини ишлаб чиқишдан иборат.

Тадқиқотнинг вазифалари:

белгилар информативлиги ва ноинформативлигининг эвристик мезонларини ишлаб чиқиш ҳамда уларнинг самарадорликларини баҳолаш;

информатив ва ноинформатив белгилар мажмуасини аниқлаш усулларини яратишда фойдаланиладиган Фишер информативлик мезонлари хусусиятларини аниқлаш;

ягона информативлик мезонидан фойдаланиб, информатив ва ноинформатив белгилар мажмуасини аниқлаш усул ва алгоритмларини яратиш;

бир жинсли нолинчи ва k -тартибли умумлашган информативлик

мезонларидан фойдаланиб информатив ҳамда ноинформатив белгилар мажмуасини аниқлаш усул ва алгоритмларини ишлаб чиқиш;

аралаш типли информативлик мезонларидан фойдаланиб информатив ва ноинформатив белгилар мажмуасини аниқлаш усул ва алгоритмларини ишлаб чиқиш;

эвристик мезонлардан фойдаланиб бошланғич белгилар тўпламини информатив, кам информатив ва ноинформатив гуруҳларга ажратиш қоидасини яратиш;

информатив ва ноинформатив белгиларни аниқлашнинг ишлаб чиқилган усулларини алгоритмик ва дастурий таъминотини яратиш.

Тадқиқотнинг объекти сифатида қийматлари “объект-хосса” жадвали кўринишда берилган объектларнинг (жараёнлар, ҳодисалар, воқеалар) белгилари (хоссалари, параметрлари, характеристикалари) олинган.

Тадқиқотнинг предмети информатив белгилар мажмуасини шакллантиришда фойдаланиладиган эвристик мезонлар, усуллар ва алгоритмлардан иборат.

Тадқиқотнинг усуллари. Ишнинг назарий тадқиқотлари математик анализ, дискрет математика ва тимсолларни аниқлаш усулларига асосланган, информатив белгилар фазосини шакллантириш ва бу фазода номаълум объектларни таниб олишнинг компьютерли моделдан тажрибалар асоси сифатида фойдаланилган ҳамда тадқиқ қилинган.

Тадқиқотнинг илмий янгилиги қуйидагилардан иборат:

белгилар информативлиги ва ноинформативлигининг эвристик мезонлари ишлаб чиқилган;

информатив ва ноинформатив белгилар мажмуасини аниқлаш усулларини яратишда фойдаланиладиган Фишер информативлик мезонларининг информативлик хусусиятлари аниқланган;

ягона информативлик мезонидан фойдаланиб, информатив ва ноинформатив белгилар мажмуасини аниқлаш усули компактлик гипотезаси асосида такомиллаштирилган ва алгоритмлари яратилган;

бир жинсли нолинчи ва k -тартибли умумлашган информативлик мезонлари асосида информатив ҳамда ноинформатив белгилар мажмуасини аниқлаш усуллари ва алгоритмлари ишлаб чиқилган;

информатив ва ноинформатив белгилар мажмуасини аниқлаш усуллари аралаш типли информативлик мезонлари асосида такомиллаштирилган ва алгоритмлари ишлаб чиқилган;

эвристик мезонлар асосида бошланғич белгилар тўпламини информатив, кам информатив ва ноинформатив гуруҳларга ажратиш қоидаси яратилган.

Тадқиқотнинг амалий натижаси қуйидагилардан иборат:

нисбатан кам сарф-ҳаражатларда олинган натижалар тўла танлов усули натижалари билан мос тушадиган информатив ва ноинформатив белгилар мажмуасини аниқлашнинг қидирув усуллари ва алгоритмлари ишлаб чиқилган;

Энг мақбул «мезон-усул» жуфтлигини танлашни сезиларли

камайтирадиган эвристик информативлик мезонининг ҳар бир кўриниши учун информатив ва ноинформатив белгилар мажмуасини аниқлашнинг мос усули ва алгоритми ишлаб чиқилган.

Тадқиқот натижаларининг ишончлилиги. Тадқиқот натижаларининг ишончлилиги белгилар фазоси ўлчамини камайтириш масаласи кўйилиши корректлиги дискрет оптимизация масаласи каби қатъий математик исботланган теорема ва леммалар кўринишидаги ҳисоблашлар, шунингдек ишлаб чиқилган усул ва алгоритмларнинг тажриба синовларида ҳамда модел ва реал масалаларда олинган натижаларни солиштириш билан изоҳланади.

Тадқиқот натижаларининг илмий ва амалий аҳамияти. Тадқиқот натижаларининг илмий аҳамияти объект, ҳодиса ва жараёнларнинг информатив таснифини шакллантиришда эвристик информативлик мезонлари асосида белгилар фазоси ўлчамини камайтиришнинг қатъий математик усулларини ишлаб чиқиш орқали тимсолларни таниб олиш назариясини ривожлантиришга хизмат қилади. Ишлаб чиқилган усул ва алгоритмлар таниб олиш ва тавсифлаш тизимлари, техник ташхислаш, робототехник тизимлар, материалшунослик, технологик жараёнлар, саноат объектларини бошқариш каби техник ва саноат иловаларни яратишда кенг кўламда қўлланилиши билан изоҳланади.

Тадқиқот натижаларининг амалий аҳамияти ишлаб чиқилган алгоритмлар ва дастурий таъминотни лойиҳалаш ва яратиладиган турли замонавий таниб олиш тизимларининг (нутқ, матн, шахсни таниб олиш ва шу кабилар) ишчи белгилар луғатини шакллантириш билан изоҳланади. Тадқиқот натижаларини кўллаш объект, ҳодиса ва жараёнларнинг информатив таснифи орқали идентификациялаш, таснифлаш ва таниб олишга сарфланадиган ресурсларни қисқартириш имконини беради.

Тадқиқот натижаларининг жорий қилиниши. Маълумотларга дастлабки ишлов бериш алгоритмлари ва усуллари асосида яратилган дастурий таъминот асосида:

датчиклардан олинган маълумотлар базасини яратиш, сақлаш, қайта ишлаш, дастлабки ишлов бериш ҳамда информатив таснифлаш алгоритмлари асосида яратилган дастурий мажмуа Навоий тоғ-металлургия комбинати марказий кон бошқармасига жорий қилинган (Ахборот технологиялари ва коммуникацияларни ривожлантириш вазирлигининг 2017 йил 21 декабрдаги 33-8/8663-сон маълумотномаси). Илмий тадқиқот натижасида КНК-270 комплексини ишлаш жараёнидаги хавфли ҳолатларни аниқлаш тезлигини ошириш асосида мазкур комплексни ишчи ҳолатда бўлмаслигини 10%га қисқартириш имконини берган;

нутқ сигналларига дастлабки ишлов бериш ва филтрлаш орқали нутқ сигналлари сифатини ошириш, информатив ва ноинформатив белгиларни танлаш алгоритмлари Ўзбекистон Республикаси ички ишлар вазирлиги тергов департаментига жорий қилинган (Ахборот технологиялари ва коммуникацияларни ривожлантириш вазирлигининг 2017 йил 21 декабрдаги 33-8/8663-сон маълумотномаси). Илмий тадқиқот натижасида нутқ сигналларини матнга ўтказиш вақтини 20%га қисқартириш имконини берган;

ходимлар ҳақидаги маълумотларга дастлабки ишлов бериш, қайта ишлаш, интеллектуал таҳлил қилиш ҳамда эвристик мезонлар асосида тегишли таклифларни шакллантириш алгоритмлари Олий ва ўрта таълим вазирлигининг “Электрон вазирлик” ахборот тизимига жорий қилинган (Ахборот технологиялари ва коммуникацияларни ривожлантириш вазирлигининг 2017 йил 21 декабрдаги 33-8/8663-сон маълумотномаси). Илмий тадқиқот натижасида кадрлар модули орқали Олий таълим муассасалари профессор-ўқитувчилари ҳақидаги маълумотларни қайта ишлашда ва тегишли таклифларни шакллантириш жараёни самарадорлигини 2 баробар ошириш имконини берган;

ишлаб чиқарилаётган маҳсулотлар маълумотлар базасини шакллантириш, дастлабки ишлов бериш ва асосий кўрсаткичларини аниқлаш дастурий воситаси “INTERVETFARM” МЧЖга жорий қилинган (Ахборот технологиялари ва коммуникацияларни ривожлантириш вазирлигининг 2017 йил 21 декабрдаги 33-8/8663-сон маълумотномаси). Илмий тадқиқот натижасида маҳсулот сифат назоратига сарфланадиган вақтни 20%га қисқартириш имконини берган;

тасвирларга дастлабки ишлов бериш, силлиқлаш, филтрлаш, турли халақитдан тозалаш орқали контур чизиқларини юқори аниқликда ифодалаш алгоритмлари “Inter sharm plast” МЧЖга жорий қилинган (Ахборот технологиялари ва коммуникацияларни ривожлантириш вазирлигининг 2017 йил 21 декабрдаги 33-8/8663-сон маълумотномаси). Илмий тадқиқот натижасида яратилган дастур контур чизиқларидан ажратиб олинган асосий (информатив) нуқталар асосида шакллантирилган махсус кодли (Gkod) файл орқали станок ишлатилганда бир метр куб хом ашёдан маҳсулот ишлаб чиқариш учун кетадиган вақт 15%га қисқартириш, хом ашёни эса ўртача 20%га тежаш ва ишлаб чиқариш самарадорлиги ўртача 1,2 марта ошириш имконини берган.

Тадқиқот натижаларининг апробацияси. Мазкур тадқиқотнинг назарий ва амалий натижалари 7 та халқаро ва 20 та республика илмий-амалий анжуманларида маъруза қилинган ва муҳокамадан ўтказилган.

Тадқиқот натижаларининг эълон қилиниши. Диссертация мавзуси бўйича жами 43 та илмий иш эълон қилинган. Жумладан, 1 та монография, Ўзбекистон Республикаси Олий аттестация комиссиясининг докторлик диссертациялари асосий илмий натижаларини чоп этиш тавсия этилган илмий нашрларда 14 та мақола, жумладан, 11 таси республика ва 3 таси хорижий журналларда нашр этилган.

Диссертациянинг ҳажми ва тузилиши. Диссертация тузилиши кириш, бешта боб, хулоса, фойдаланилган адабиётлар рўйхати ва иловадан иборат. Диссертациянинг ҳажми 175 бетни ташкил этган.

ДИССЕРТАЦИЯНИНГ АСОСИЙ МАЗМУНИ

Кириш қисмида ўтказилган тадқиқотларнинг долзарблиги асослаб берилган, тадқиқотнинг мақсад ва вазифалари, объекти ва предмети белгилаб берилган. Ўзбекистон Республикаси фан ва технологияларининг устувор йўналишларига тадқиқотнинг мослиги кўрсатилган, тадқиқотнинг илмий янгилиги ва амалий натижалари кўрсатиб ўтилган, олинган натижаларнинг назарий ва амалий аҳамияти очиб берилган, тадқиқот натижаларининг амалиётга татбиқи, ишнинг синов натижалари, нашр этилган ишлар ва диссертация тузилмаси тўғрисидаги маълумотлар келтирилган.

Диссертациянинг **“Маълумотларга дастлабки ишлов бериш муаммолари”** деб номланган биринчи бобда дастлабки маълумотлардаги турли хил бўшлиқлар, халақит ва сакрашларни бартараф этиш орқали ушбу маълумотлар сифатини яхшилаш ҳамда ўзгарувчилар типларини унификациялаш, яъни турли типли шкалаларда ўлчанадиган ўзгарувчиларни ягона типли шкалада ўлчашга келтириш масалалари таҳлил қилинган.

Бўшлиқлар, халақит ёки сакрашларга эга бўлган маълумотлар паст сифатли маълумотлардир.

Маълумотлар тўлиқ бўлмаганда асосий статистик характеристикалар (масалан, математик кутилма ёки дисперсия)нинг ўзгариши бўшлиқлар сонига тўғри пропорционал ўсганлиги сабаб, уларга математик усулларни тадбиқ этиш мураккаблашади.

Бугунги кунда ноаниқ маълумотлардаги муаммоларни ҳал этишнинг кўплаб усуллари мавжуд. Диссертация ишида кенг тарқалган усуллар батафсил таҳлил этилган, жумладан, дастлабки танланмадан тўлиқ бўлмаган объектларни чиқариб ташлаш усули, устуннинг ўртача қиймати билан бўшлиқларни тўлдириш усули, сплайн-интерполяция усули, ZET алгоритми, RESAMPLING усули ва бошқалар.

Ушбу бобда сакраш ва халақитли маълумотларни таҳлил қилиш муаммоларини ҳал қилишда фойдаланиладиган силлиқлаш ва филтрлашнинг энг кенг тарқалган усуллари кўриб чиқилган.

Агар тажрибавий кузатиш натижалари турли типли белгилар қийматлари билан ифодаланган бўлса, у ҳолда бундай маълумотлар турли типли дейилади.

Турли типли кўрсаткичларнинг биргаликдаги таҳлил қилиш имкониятини мавжуд эмаслиги тажрибавий маълумотларни айрим синфлари умуман қайта ишланмаслиги ва турли типли белгиларнинг ажраган таҳлилидан олинган натижалар эса тадқиқ қилинаётган ҳодисанинг ички боғланишлари хилма-хиллигини акс эттира олмаслиги мумкин ва шунинг учун бундай ёндошувлар амалий қизиқиш касб этмайди.

Кўрсатилган муаммоларни ечишнинг муҳим ёндошувларидан бири бу номиқдор ўзгарувчиларни рақамлаштириш ғоясига асосланган. Мазкур ғоянинг мазмуни сифат ва номинал белгилар градациясига ҳақиқий сонли қийматларни (сонли меткаларни) қўшиб ёзишдан иборат. Бундай ёндошувнинг афзаллиги турли типли маълумотларга сонли ахборотларни

таҳлил қилишнинг кўплаб математик усуллари тадбиқ этиш имконини беришидадир.

Дастлабки ишлов бериш босқичида маълумотлар сифатини яхшилаш ва ўзгарувчилар типини бир хиллаштириш масаласини ечиш билан бир қаторда ноинформатив белгиларни чиқариб ташлаш ҳисобига тадқиқ қилинаётган объект тавсифи фазоси ўлчамини камайтириш зарурияти пайдо бўлади. Бу зарурият айниқса таснифлаш масалалари учун муҳим, чунки айнан бир хил объект танланмасининг ифодаланиш даражаси белгилар фазоси ўлчамига тесқари пропорционалдир. Кичик ҳажмли ўқув танланма ҳолатида ноинформатив белгиларнинг мавжудлиги маълумотларни қайта ишлаш сифатини ёмонлаштириши мумкин.

1.2-параграфда келтирилган таснифлаш масалаларида белгилар фазоси ўлчамини камайтириш муаммосини шакллантириш шуни кўрсатадики, маълумотлар таҳлили масалаларида, хусусан таснифлаш масаласида дастлабки белгилар фазоси ўлчамини камайтириш моҳиятига кўра дастлабки белгилар $x = (x^1, x^2, \dots, x^N)$ тизимдан аввалгисига қараганда камроқ ($\ell < N$) белгиларни ўз ичига олган янги $z = (z^1, z^2, \dots, z^\ell)$ тизимга ўтишдир.

Одатда оптималлаштириш масаласини ечиш орқали янги белгилар дастлабки белгилардан функция кўринишидаги янги белгилар шакллантирилади, яъни $z = F(x)$.

Бунда шундай z белгилар тизимини топиш керакки, бу белгилар тизими учун қуйидаги тенглик бажарилсин

$$I(\tilde{z}) = \max_{F \in \Omega} \{I(z)\} \quad (1)$$

бу ерда $I(z)$ - ℓ -ўлчовли z белгилар тизимининг берилган информативлик ўлчови, F -дастлабки x^1, x^2, \dots, x^N белгиларни мумкин бўлган алмаштиришлар синфи, у умумий ҳолда ночизиқли, дискрет, узлуксиз ёки мантиқий алмаштириш турларидан бири орқали тақдим этилиши мумкин. $I(z)$ -информативлик ўлчови ва мумкин бўлган F алмаштиришлар синфини аниқ танлови ўлчамни камайтиришнинг аниқ бир усулига олиб келади.

Шундай қилиб, дастлабки белгилар фазоси ўлчамини камайтиришни N -ўлчовли x векторни ℓ -ўлчовли z векторга акслантириш каби талқин қилиш мумкин, буни умумий ҳолда $z = F(x)$ кўринишида ифодалаш мумкин, бунда F -мумкин бўлган алмаштириш, $\ell \leq N$.

Одатда дастлабки белгилар фазосида алмаштиришнинг икки усулидан фойдаланилади.

Алмаштиришнинг биринчи усулини қуйидаги функциялар системаси кўринишида ифодалаш мумкин

$$\begin{cases} z^1 = f_1(x^1, x^2, \dots, x^N), \\ z^2 = f_2(x^1, x^2, \dots, x^N), \\ \dots, \\ z^\ell = f_\ell(x^1, x^2, \dots, x^N), \end{cases} \quad (2)$$

бу ерда z^1, z^2, \dots, z^ℓ -янги белгилар, x^1, x^2, \dots, x^N -дастлабки белгилар функциялари

кўринишида ифодаланади (одатда бу функция чизиқли) ва бунда $\ell < N$ бўлади.

Иккинчи усулни қуйидаги функциялар орқали ифодалаш мумкин

$$z^i = f_i(x^i); i = \overline{1, N}, \quad (3)$$

$$f(x^i) = \begin{cases} 0, & \text{агар } x^i \text{ белгичикариб юборилса;} \\ 1, & \text{агар } x^i \text{ белгиколдирилса.} \end{cases}$$

Ушбу ҳолда белгиларни янги системаси дастлабки белгилар тўпламининг қисм тўплами сифатида шакллантирилади.

(2) ва (3) кўринишдаги алмаштиришлардан фойдаланишга асосланган усуллар дастлабки белгилар фазоси ўлчамини камайтириш усуллари деб аталади.

Шуни таъкидлаб ўтиш керакки, диссертация ишида дастлабки белгилар фазоси ўлчамини камайтириш масаласини ечишга йўналтирилган усул ва алгоритмлар (3) кўринишидаги алмаштиришга асосланиб ишлаб чиқилган.

Диссертациянинг биринчи боби белгилар фазоси ўлчамини камайтириш муаммосини ечишда фойдаланиладиган эвристик мезонлар ва усуллар таҳлили билан яқунланади. Ушбу натижалар асосида диссертациянинг мақсади ва вазифалари шакллантирилган.

Диссертациянинг **“Белгилар информативлигининг эвристик мезонлари ва уларнинг самарадорлиги”** деб номланган иккинчи бобида белгилар информативлигининг эвристик мезонлари хусусий ва умумий кўринишлари баёни келтирилган. Бундан ташқари Фишер информативлик мезонини тадқиқ қилиш орқали унинг хусусиятлари ва самарадорлигини баҳолаш кўрсатиб берилган.

2.1-параграфда асосий тушунчалар, белгилашлар ва зарурий таърифлар келтирилган.

Фараз қилайлик, $x_{11}, x_{12}, \dots, x_{1m_1}, x_{21}, x_{22}, \dots, x_{2m_2}, \dots, x_{p1}, x_{p2}, \dots, x_{pm_p}$ ўқув танланмаси берилган бўлиб, бунда ҳар бир $x_{p1}, x_{p2}, \dots, x_{pm_p}$ объектлар тўплами маълум бир $X_p, p = \overline{1, r}$ синфга тегишли бўлсин. Ҳар бир x_{pi} -объектнинг белгилари N -ўлчовли сон қийматли вектор кўринишида ифодаланган бўлсин, яъни $x_{pi} = (x_{p1}, x_{p2}, \dots, x_{pm_p})$.

Берилган ўқув танланмалари $x_{p1}, x_{p2}, \dots, x_{pm_p} \in X_p, p = \overline{1, r}$ учун информатив белгилар қисм фазосини бир қийматли характерловчи $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_N), \lambda_i \in \{0, 1\}, i = \overline{1, N}$ векторни киритиб оламиз. Бу ерда x_{pi} - N -ўлчовли белгилар фазоси вектори. λ векторнинг бирга тенг бўлган компоненталарига мос келувчи белгилар ажратиб олинаётган қисм фазода иштирок этишини билдирса, нолга тенг бўлган компоненталари эса унга мос белгиларни ажратиб олинаётган қисм фазода иштирок этмаслигини билдиради.

Белгилар фазоси $\{x = (x^1, x^2, \dots, x^N)\}$ ни Евклид фазоси деб ҳисоблаймиз ва R^N орқали белгилаймиз.

1-таъриф. $R^N|_{\lambda} = \{x|_{\lambda} = (\lambda_1 x^1, \lambda_2 x^2, \dots, \lambda_N x^N)\}$ фазо $R^N = \{x = (x^1, x^2, \dots, x^N)\}$ фазонинг λ бўйича акслантирмаси дейилади.

R^N фазодаги x, y объектлар орасидаги акслантирма масофа сифатида $R^N|_{\lambda}$ фазодаги икки $x|_{\lambda}$ ва $y|_{\lambda}$ объектлар орасидаги $\|x - y\|_{\lambda} = \sqrt{\sum_{i=1}^N \lambda_i (x^i - y^i)^2}$

Евклид масофасини оламыз.

2-таъриф. λ вектор ℓ -информатив дейилади, агарда унинг компоненталари йиғиндиси ℓ га тенг бўлса, яъни $\sum_{i=1}^N \lambda_i = \ell$.

Ҳар бир қисм системадаги λ ℓ -информатив вектор учун унга мос ℓ -ўлчовли белгилар қисм фазоси аниқланган ва бу қисм фазоларда λ мос $\|x\|_{\lambda} = \sqrt{\sum_{j=1}^N \lambda_j (x^j)^2}$ Евклид нормасини киритамиз. X_p синфнинг характерловчи

ўртача объекти \bar{x}_p қуйидагича аниқланади: $\bar{x}_p = \frac{1}{m_p} \sum_{i=1}^{m_p} x_{pi}, p = \overline{1, r}$.

Қуйидаги функцияни киритиб оламыз:

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

$S_p(\lambda)$ функция X_p синфдаги λ вектор асосида ажратиб олинган объектларнинг ўртача тарқоқлигини ифодалайди. Информативлик мезони сифатида қуйидаги функционални оламыз:

$$I_1(\lambda) = \frac{\sum_{p,q=1}^r \|\bar{x}_p - \bar{x}_q\|_{\lambda}^2}{\sum_{p=1}^r S_p^2(\lambda)}. \quad (4)$$

(4) функционал Фишер функционалининг содда кўринишини ифодалайди.

$a = (a^1, a^2, \dots, a^N)$ ва $b = (b^1, b^2, \dots, b^N)$ белгилашларни киритиб оламыз, бунда ерда $a^j = \sum_{p,q=1}^r (\bar{x}_p^j - \bar{x}_q^j)^2$, $b^j = \sum_{p=1}^r \frac{1}{m_p} \sum_{i=1}^{m_p} (\bar{x}_{pi}^j - \bar{x}_{qi}^j)^2$, $j = \overline{1, N}$;

У ҳолда (4) функционал (5) кўринишга келади. Бу кўринишдаги мезонни $I(\lambda)$ орқали белгилаб оламыз

$$I(\lambda) = \frac{(a, \lambda)}{(b, \lambda)}, \quad (5)$$

бу ерда $(*, *)$ -векторларнинг скаляр кўпайтмаси.

a^j, b^j коэффициентлар λ га боғлиқ бўлмагани учун олдиндан ҳисоблаб олиш мумкин. $I(\lambda)$ мезонни ҳар бир λ га мос қийматини топиш учун N тартибли амаллар бажариш зарур бўлади.

3-таъриф. $\forall \alpha \in R, (\alpha \neq 0)$ учун $f(\alpha x_1, \alpha x_2, \dots, \alpha x_k) = \alpha^k f(x_1, x_2, \dots, x_k)$ тенглик бажарилса, f k -тартибли бир жинсли функция дейилади.

Мазкур таърифни инобатга олиб, 2.1-параграфда информатив белгилар

фазосини шакллантиришнинг 0-тартибли бир жинсли умумий мезони қуйидагича шакллантирилади:

$$I(\lambda) = \frac{(a, \lambda)^r}{\prod_{j=1}^r (b^{(j)}, \lambda)}; \quad (6)$$

$$I(\lambda) = \prod_{j=1}^r \frac{(a^{(j)}, \lambda)}{(b^{(j)}, \lambda)}. \quad (7)$$

(6) ва (7) мезонлардан таснифланаётган синфлар сони икки ёки ундан кўп бўлган ҳолларда фойдаланилса мақсадга мувофиқ бўлади. Бундай мезонлардан ташқари k -тартибли бир жинсли ва бир жинсли бўлмаган мезонлар ҳам мавжуд. Ушбу мезонлар кам тадқиқ этилганлиги боис, ҳозирги кунда бундай мезонлар асосида информатив белгилар мажмуасини аниқлашнинг усул ва алгоритмлари ишлаб чиқилмаган.

Юқорида келтирилган белгилаш ва тушунчалардан фойдаланган ҳолда, қуйидаги кўринишда берилган мезонни кўриб чиқамиз:

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

бу ерда c қандайдир ташқи таъсир қилувчи омиллар вектори. Масалан, мазкур векторнинг ҳар бир компонентаси белгиларни ўлчаш учун талаб қилинадиган воситаларни ифодаланиши мумкин.

(8) функционал Фишер типдаги функционални “-1”-тартибли бир жинсли мезон кўринишини ифодалайди. Информатив белгиларни танлашда (8) функционалнинг максимизация масаласи ечилади.

Агар синфлар сони учта ($r=3$) деб олсак, у ҳолда қуйидаги кўринишдаги функционалга эга бўламиз:

$$I(\lambda) = \frac{(a, \lambda)}{(b, \lambda)(c, \lambda)(d, \lambda)}. \quad (9)$$

Умумий ҳолда эса k -тартибли бир жинсли мезонни қуйидаги икки кўринишда ифодалаш мумкин:

$$I(\lambda) = \frac{\prod_{j=1}^r (a^{(j)}, \lambda)}{(b, \lambda)}, \quad (10)$$

$$I(\lambda) = \frac{(a, \lambda)}{\prod_{i=1}^r (b^{(i)}, \lambda)}. \quad (11)$$

(10) мусбат даражали k -тартибли бир жинсли мезон, (11) эса манфий даражали k -тартибли бир жинсли мезондир. k -тартибли бир жинсли мезоннинг умумий кўриниши қуйидагича бўлади:

$$I(\lambda) = \frac{\prod_{j=1}^t (a^{(j)}, \lambda)}{\prod_{i=1}^h (b^{(i)}, \lambda)}, (t, h \in N). \quad (12)$$

Шуни алоҳида таъкидлаб ўтиш жоизки, k -тартибли бир жинсли мезонлар ҳар бир синф ички жойлашуви қонуниятларини алоҳида эътиборга

олиши ва синфлараро масофаларни ўзгартирмаслиги билан 0-тартибли бир жинсли мезонлардан устун ҳисобланади.

Бир жинсли мезонлардан ташқари амалий масалаларни ечишда бир жинсли бўлмаган мезонлардан ҳам фойдаланилади, масалан,

$$I(\lambda) = \frac{(a, \lambda)}{(b, \lambda)} + \frac{(a^*, \lambda)}{(b^*, \lambda)(c^*, \lambda)}. \quad (13)$$

(13) мезон 0-тартибли бир жинсли ва k -тартибли бир жинсли мезонларни йиғиндилари кўринишида эканлигини пайқаш қийин эмас.

Умумий ҳолда Фишер типигаги бир жинсли бўлмаган мезонлар (6)-(8), (10)-(12) мезонларнинг комбинацияларидан ташкил топади.

Диссертация ишининг 2.2-параграфида (5) функционал кўринишида берилган информативлик мезонларининг айрим хусусиятлари келтирилган.

Λ^ℓ -тўпдам барча ℓ -информатив векторлар тўплари бўлсин. ℓ -информатив вектор таърифига кўра, Λ^ℓ тўпдамнинг қуввати C_N^ℓ га тенг.

Қуйидаги функцияни киритиб оламиз

$$F(\ell) = \max_{\lambda \in \Lambda^\ell} I(\lambda). \quad (14)$$

Берилган ℓ да $I(\lambda_\ell) = \max_{\lambda \in \Lambda^\ell} I(\lambda)$ тенглик ўринли бўладиган $\lambda_\ell \in \Lambda^\ell$ ℓ -информатив вектор топилсин ва $F(\ell)$ функциянинг аналитик ифодаси аниқлансин.

Ушбу масалани ечиш келтирилган 1-6-леммалар, шунингдек қуйидаги теоремалар орқали амалга оширилади.

Ҳақиқий a, b ва $c \geq 0, d > 0$ ($a+c \geq 0, b+d > 0$) сонлар берилган бўлсин. У ҳолда қуйидаги леммалардан бири ўринли бўлади:

1-лемма. Агар $\begin{cases} a > 0 \\ b > 0 \end{cases}$ ва $\frac{c}{d} > \frac{a}{b}$ бўлса, у ҳолда $\frac{a}{b} < \frac{a+c}{b+d} < \frac{c}{d}$ муносабат ўринли.

2-лемма. Агар $\begin{cases} a > 0 \\ b > 0 \end{cases}$ ва $\frac{c}{d} < \frac{a}{b}$ бўлса, у ҳолда $\frac{a}{b} > \frac{a+c}{b+d} > \frac{c}{d}$ муносабат ўринли.

3-лемма. Агар $\begin{cases} a < 0 \\ b < 0 \end{cases}$ ва $\frac{c}{d} < \frac{a}{b}$ бўлса, у ҳолда $\frac{a}{b} > \frac{a+c}{b+d} < \frac{c}{d}$ муносабат ўринли.

4-лемма. Агар $\begin{cases} a < 0 \\ b < 0 \end{cases}$ ва $\frac{c}{d} > \frac{a}{b}$ бўлса, у ҳолда $\frac{a}{b} < \frac{a+c}{b+d} > \frac{c}{d}$ муносабат ўринли.

5-лемма. Агар $\begin{cases} a \geq 0 \\ b \leq 0 \end{cases}$ бўлса, у ҳолда $\frac{a+c}{b+d} \geq \frac{c}{d}$ муносабат ўринли.

6-лемма. Агар $\begin{cases} a \leq 0 \\ b \geq 0 \end{cases}$ бўлса, у ҳолда $\frac{a+c}{b+d} \leq \frac{c}{d}$ муносабат ўринли.

1-теорема. $\forall \ell (\ell > 0)$ учун $\max \{I(\lambda) : \lambda \in \Lambda^\ell\} = \frac{a^1 + \sum_{k=1}^{\ell-1} a^{j_k}}{b^1 + \sum_{k=1}^{\ell-1} b^{j_k}}$ бўлади.

(14) функция учун қуйидаги теорема ўринли.

2-теорема. $F(\ell)$ функция қиймати ℓ аргумент ортганда ўсмайди.

Фишер типигаги (5) информативлик мезонлари хоссаларини тадқиқи олинган натижаларни информатив белгилар тўпламини танлаш усулини

яратишда фойдаланиш мумкин.

Диссертациянинг “Белгилар информативлигининг эвристик мезонлари асосида объектларни информатив тавсифлаш усуллари” деб номланган учинчи бобида тимсолларни таниб олиш масаласининг информатив белгилар мажмуасини аниқлаш усул ва алгоритмлари таклиф этилган.

3.1-параграфда (5) функционал кўринишидаги информативлик мезони асосида информатив белгилар мажмуасини аниқлашнинг регуляр, рекуррент, қисмий танлов, тартиблаш ва "Дельта" усуллари таклиф этилган. Бунда информатив белгилар мажмуасини аниқлашнинг оптимизацион масаласи куйидагича:

$$\begin{cases} I(\lambda) \rightarrow \max; \\ \lambda \in \Lambda^\ell, \end{cases} \quad (15)$$

бу ерда Λ^ℓ -тўплам барча ℓ -информатив белгилар тўплами.

Регуляр усулида ушбу масалани ечиш учун куйидаги вектор-функциядан фойдаланилган:

$$\varphi(\lambda) = a(b, \lambda) - b(a, \lambda), \quad (16)$$

бу вектор-функция $I(\lambda)$ функциянинг λ нуқтага нисбатан тезкор ўсиш йўналишини кўрсатиб беради. Мазкур усул куйидаги икки теорема ва хоссага асосланган.

Фараз қилайлик $\lambda \in \Lambda^\ell$ вектор танлаган бўлсин. $\forall \mu \in \Lambda^\ell$ вектор учун куйидаги теорема ўринли.

3-теорема. $I(\lambda) > I(\mu)$ бўлиши учун $(\varphi(\lambda), \mu) > 0$ шартни қаноатлантирувчи μ векторнинг мавжуд бўлмаслиги зарур ва етарли.

1-хосса. $\forall \lambda (\lambda \in \Lambda^\ell)$ учун $(\varphi(\lambda), \mu(\lambda)) \geq 0$ тенгсизлик ўринли.

3-теорема ва 1-хоссадан $I(\lambda) \leq I(\mu(\lambda))$ эканлиги келиб чиқади.

4-теорема. Агар $I(\lambda) = I(\mu(\lambda))$ бўлса, у ҳолда $I(\lambda) = \max\{I(\eta) | \eta \in \Lambda^\ell\}$ бўлади.

Олинган ечим оптималлигини мазкур теорема таъминлаб беради, яъни $I(\lambda)$ функционал Λ^ℓ тўпламининг λ элементида ўзининг максимумига эришади.

Регуляр усули 3- ва 4-теоремаларга асосланган бўлиб, у (5) функционални максималлаштиришга итератив процедура шаклида жорий этилган. Дастлабки кадамда ихтиёрий ℓ -информатив λ вектор олинади.

Масалан, $\lambda = (\overbrace{1, 1, \dots, 1}^{\ell}, 0, 0, \dots, 0)$. Кейинги ҳар бир кадамда янги вектор λ олдингисига $\mu(\lambda)$ операторни кўллаш орқали ҳосил қилинади, яъни $\lambda = \mu(\lambda)$. Итерацион жараён $I(\lambda)$ функционал ўсишдан тўхтагунича давом этади. Агар ўсиш жараёни тўхтаса, яъни $I(\lambda) = I(\mu(\lambda))$ шарт бажарилса, у ҳолда λ -оптимал ечим. Одатда 3-4 кадамда оптимал ечим олишга эришилади.

Таклиф этилган қисмий танлов усулининг мақсади $I(\lambda)$ функционални максимумга эриштирувчи λ ℓ -информатив векторни топишдан иборат. Бунда берилган $\ell \geq 2$ учун $I(\lambda)$ функционалнинг ҳисоблашлар сони куйидагига тенг:

$$\frac{\ell \prod_{i=1}^{\ell-1} (N - \ell + i)}{\ell!}.$$

Тадқиқот ишида таклиф этилган информатив белгилар мажмуасини аниқлашнинг тартибланиш усули (15) масалани ечишнинг ғояси қуйидагича.

Фараз қилайлик, a ва b векторлар компоненталарига нисбатан мос равишда қуйидагича тартибланган бўлсин:

$$\frac{a_1}{b_1} \geq \frac{a_2}{b_2} \geq \dots \geq \frac{a_N}{b_N}. \quad (17)$$

У ҳолда $\lambda = (\underbrace{1,1,\dots,1}_{lma}, \underbrace{0,0,\dots,0}_{N-lma})$ вектор тартибланиш усули натижасига кўра

(15) масаланинг оптимал ечими ҳисобланади. Бироқ тартибланиш усули ҳар доим ҳам Фишер мезонига нисбатан оптимал ечимни таъминламайди. Масалан, $a = (5,10,10,1)$, $b = (1,50,50,19)$ ва $N = 4$, $\ell = 2$ бўлганда оптимал ечим $\lambda = (1,1,0,0)$ вектор бўлмай, балки $\lambda = (1,0,0,1)$ вектордир.

Қуйида тартибланиш усули (15) масаланинг оптимал ечимини берувчи шартлари келтирилган.

Қуйидаги белгиланишларни киритайлик:

$$A = \sum_{i=1}^l a_i, B = \sum_{i=1}^l b_i, \begin{cases} \Delta a_{ij} = a_j - a_i \\ \Delta b_{ij} = b_j - b_i, i = \overline{1, l}, j = \overline{l+1, N} \end{cases}, \lambda^0 = \left(\underbrace{1,1,\dots,1}_{lma}, \underbrace{0,0,\dots,0}_{N-lma} \right).$$

Юқорида келтирилган леммаларда $a = \Delta a_{ij}$, $b = \Delta b_{ij}$, $c = A$, $d = B$ деб олсак,

у ҳолда $\forall i, j (i = \overline{1, \ell}, j = \overline{\ell+1, N})$ лар учун $\begin{cases} A + \Delta a_{ij} \geq 0, \\ B + \Delta b_{ij} > 0 \end{cases}$ бўлиб, 1-6-леммалардан

бири ўринли бўлади.

5-теорема. Тартибланган (17) кетма-кетлик ёрдамида танлаб олинган $\lambda^0 = (\underbrace{1,1,\dots,1}_{lma}, \underbrace{0,0,\dots,0}_{N-lma})$ вектор (15) масаланинг оптимал ечим бўлиши учун 2- ва 4-леммалар шартларини қаноатлантирувчи $a = \Delta a_{ij}$, $b = \Delta b_{ij}$ ларнинг мавжуд бўлмаслиги зарур ва етарли.

Тартибланиш усули асосида ишлаб чиқилган «Дельта» усули қуйидагича амалга оширилади.

Фараз қилайлик, $\lambda = (\underbrace{1,1,\dots,1}_{lma}, \underbrace{0,0,\dots,0}_{N-lma})$ вектор тартибланиш усули орқали

олинган ечим ва у (15) масаланинг оптимал ечими бўлмасин. У ҳолда (15) масаланинг оптимал ечимини аниқлаш учун 2- ва 4-леммалар асосида алмаштиришлар бажарилади. Алмаштириш жараёни 2- ва 4-леммалар шартларини қаноатлантирувчи Δa_{ij} ва Δb_{ij} мавжуд бўлмаслик шarti бажарилмагунча давом эттирилади. Агар Δa_{ij} ва Δb_{ij} 2- ва 4-леммалар шартларини қаноатлантирувчи Δa_{ij} ва Δb_{ij} қолмаса, 5-теорема натижасига кўра ҳосил қилинган ечим оптимал.

Кўп ҳолларда дастлаб танлаб олинган λ вектор (15) масаланинг оптимал ечимини таъминлайди, шунинг учун қуйида келтириладиган теорема масаланинг қачон оптимал ечим эканлигини аниқлаш имкони

беради.

$\forall \lambda \in \Lambda^l$ танланган бўлсин.

6-теорема. Танланган λ вектор (15) масаланинг оптимал ечими бўлиши учун 2-, 4- ва 5-лемма шартларини қаноатлантирувчи $a = \Delta a_{ij}$ ва $b = \Delta b_{ij}$ ($i = \overline{1, l}, j = \overline{l+1, N}$) ларнинг мавжуд бўлмаслиги зарур ва етарли.

«Дельта» усулида функционалнинг қиймати ва λ векторнинг компоненталари тартиблаш усулидаги каби шакллантирилади.

Диссертация ишининг 3.2-параграфиди (15) масаланинг (6) функционал кўринишида берилган мезон асосида ечими кўриб чиқилган.

Дастлаб λ векторни $\lambda^0 = (\underbrace{1, 1, \dots, 1}_{\ell \text{ ма}}, \underbrace{0, 0, \dots, 0}_{N-\ell \text{ ма}})$ кўринишида оламиз ва қуйидаги

белгилашларни киритамиз:

$$A = \sum_{i=1}^l a_i, \quad B_j = \sum_{i=1}^l b_i^j, \quad j = \overline{1, r},$$

шунингдек, ёрдамчи C векторни киритиб оламиз ва унинг компоненталари

$$c_i = \frac{a_i}{A} - \sum_{j=1}^r \frac{b_i^{(j)}}{B_j} \text{ формула орқали ҳисобланади.}$$

$\forall \mu \in \Lambda^\ell$ ва λ^0 вектор асосида шакллантирилган C векторлар берилган бўлсин.

7-теорема. Танлаб олинган λ вектор (15) масаланинг оптимал ечими бўлиши учун $(C, \mu) > 0$ шартни қаноатлантирувчи μ векторнинг мавжуд бўлмаслиги зарур ва етарли.

Ушбу теоремага асосланган усул диссертация ишида “Градиент-1” деб номланади ва бунда μ вектор регуляр усулидаги каби шакллантирилади.

Фараз қилайлик, (8) мезон берилган бўлсин. У «-1» тартибли бир жинсли мезон бўлиб, унга мос (15) масалани кўрайлик.

Қуйидаги белгилашларни киритайлик:

$$A = \sum_{i=1}^l a_i, \quad B = \sum_{i=1}^l b_i, \quad C = \sum_{i=1}^l c_i, \quad \begin{cases} \Delta a_{ij} = a_j - a_i, \\ \Delta b_{ij} = b_j - b_i, i = \overline{1, l}, j = \overline{l+1, N}. \end{cases}$$

$\forall a, b, c$ ва $d, e, f > 0$ ($b+e > 0, c+f > 0$) ҳақиқий сонлар учун қуйидаги

леммалар ўринли:

7-лемма. Агар $\frac{a}{d} \geq \frac{b}{e} + \frac{c}{f} + \frac{bc}{ef}$ бўлса, у ҳолда $\frac{d}{ef} \leq \frac{d+a}{(b+e)(c+f)}$ бўлади.

8-лемма. Агар $\frac{a}{d} < \frac{b}{e} + \frac{c}{f} + \frac{bc}{ef}$ бўлса, у ҳолда $\frac{d}{ef} > \frac{d+a}{(b+e)(c+f)}$ бўлади.

Фараз қилайлик, $\forall \lambda \in \Lambda^\ell$ танланган бўлсин.

8-теорема. Танланган λ вектор (15) масаланинг оптимал ечими бўлиши учун 7-лемма шартларини қаноатлантирувчи $a = \Delta a_{ij}, b = \Delta b_{ij}$ ва $c = \Delta c_{ij}$ ($i = \overline{1, l}, j = \overline{l+1, N}$) ларнинг мавжуд бўлмаслиги зарур ва етарли.

Мазкур усулда функционалнинг қиймати ва λ векторнинг компоненталари тартиблаш усулидаги каби шакллантирилади.

Бундан ташқари, диссертация ишининг 3.2-параграфиди (10) кўринишдаги мусбат тартибли умумий мезон асосида информатив белгиларни ажратишнинг (15) масаласи кўриб чиқилган.

Ҳисоблашларни соддалаштириш мақсадида қуйидаги белгилашларни киритиб оламиз:

$$A^{(j)} = (a^{(j)}, \lambda), B = (b, \lambda), j = \overline{1, r}, A_1^{(j)} = (a^{(j)}, \mu), B_1 = (b, \mu).$$

шунингдек, (15) масалани ечиш учун $C = (c_1, c_2, \dots, c_N)$ оптимал ечимга йўналтирувчи векторни киритиб оламиз ва уни компоненталари

$$c_i = \frac{\sum_{t=1}^r a_i^{(t)}}{w^{r-1}} - I_\lambda b_i, i = \overline{1, N} \text{ формула орқали аниқланади:}$$

Фараз қилайлик, $\lambda, \mu \in \mathbb{R}^l$ - информатив вектор ва $b_i > 0, i = \overline{1, N}$ бўлсин. У ҳолда $I(\lambda), I(\mu)$ ва шакллантирилган C вектор учун қуйидаги теорема ўринли.

9-теорема. $I(\lambda) < I(\mu)$ бўлиши учун $(C, \mu) > 0$ бўлиши етарли.

9-теоремага асосланган усул "Градиент-2" деб номланади. Ушбу усулда μ вектор регуляр усулидаги каби шакллантирилади. Бунда мазкур усулдан манфий $1-r$ даражали бир жинсли мезонлар асосида информатив белгилар тўпламини аниқлашда фойдаланиш мумкин.

Функционал (11) Фишер типидagi манфий даражали $k (k \leq 0)$ бир жинсли мезоннинг умумий кўринишининг ифодалайди.

(15) масалани ечиш учун $C = (c_1, c_2, \dots, c_N)$ градиент векторни киритиб оламиз. Бу вектор компонентлари $c_i = \frac{a_i}{w^{r-1}} - I_\lambda \sum_{j=1}^r b_i^{(j)}, i = \overline{1, N}$ формула ёрдамида аниқланади. $C = (c_1, c_2, \dots, c_N)$ вектор учун 9-теорема ўринли.

Фишер типидagi мезонлар комбинацияларидан ташкил топган мезонлар асосида информатив белгилар фазосини шакллантириш усули бу мезонларга мос йўналтирувчи векторлар комбинациялари асосида шакллантирилади. Бир жинсли бўлмаган мезонларнинг комбинациясидан ташкил топган C вектор учун 9-теоремалар ўринли эканлиги диссертация ишида кўрсатилган.

Диссертациянинг **“Ноинформатив белгилар мажмуасини аниқлаш усуллари”** деб номланган тўртинчи бобиди таклиф этилган эвристик информативлик мезонлари асосида ноинформатив ва кам информатив белгилар мажмуасини аниқлаш усуллари баён этилган.

$R^N \setminus \Lambda^\ell$ тўплам ноинформатив белгилар тўплами деб аталади ва уни \aleph орқали белгиланади, яъни $\aleph = R^N \setminus \Lambda^\ell$. Бунда бошланғич белгилар фазоси белгилар фазоси $\{\Lambda^\ell, \aleph\}$ кўринишда ифодаланади, яъни $R^N = \Lambda^\ell \cup \aleph$.

Ҳозирги кунда мавжуд усулларнинг катта қисми информатив белгиларни аниқлашга йўналтирилган, бироқ ноинформатив белгиларни чиқариш орқали информатив белгиларни аниқловчи усуллар ҳам мавжуд.

Фараз қилайлик, Λ^ℓ информатив белгиларни аниқлаш усулларида олинган информатив белгилар тўплами ва \aleph эса ноинформатив белгиларни аниқлаш усулларида олинган ноинформатив белгилар тўплами бўлсин. У ҳолда:

1. Агар $\Lambda^\ell \cap \aleph = \emptyset$ бўлса, у ҳолда $\Lambda^\ell \cup \aleph = R^N$ бўлиб, $\dim(\Lambda^\ell) + \dim(\aleph) = N$ бўлади.

2. Агар $K = \Lambda^\ell \cap \aleph \neq \emptyset$ бўлса, у ҳолда $\Lambda^\ell \cup \aleph \subset R^N$ бўлиб, $\Lambda^\ell \setminus K$ информатив белгилар тўплами, $\aleph \setminus K$ ноинформатив белгилар тўплами, K эса кам информатив белгилар тўплами дейилади. Бунда $\dim(\Lambda^\ell) + \dim(\aleph) > N$ бўлиб, $\dim(\Lambda^\ell \setminus K) + \dim(\aleph \setminus K) + \dim(K) = N$ бўлади ва бошланғич белгилар фазоси белгилар фазоси $\{\Lambda^\ell, K, \aleph\}$ кўринишда ифодаланади, яъни $R^N = \Lambda^\ell \cup K \cup \aleph$.

Диссертация ишида таклиф этилган ноинформатив белгилар мажмуасини аниқлаш усуллари 3.1-параграфда келтирилган информатив белгилар мажмуасини аниқлаш усуллари модификацияси ҳисобланади.

Ноинформатив белгилар мажмуасини аниқлашнинг таклиф этилаётган модификацион регуляр усули (5) кўринишга келтириш мумкин бўлган информативлик мезонларига асосланиб ва ҳал этилиши лозим бўлган масала қуйидагича шакллантирилади

$$\begin{cases} I(\lambda) \rightarrow \min; \\ \eta \in \aleph^\wp, \end{cases} \quad (18)$$

бу ерда \aleph^\wp - \wp -ноинформатив барча белгилар тўплами.

(18) масалани ечиш учун η нуктада $I(\eta)$ функционални тезкор камайиш йўналишини кўрсатиб берувчи қуйидаги вектор функцияни киритамиз: $\varphi(\eta) = a(b, \eta) - b(a, \eta)$,

10-теорема. $I(\eta) < I(\mu)$ бўлиши учун $(\varphi(\eta), \mu) < 0$ шартни қаноатлантирувчи μ \wp -ноинформатив векторнинг мавжуд бўлмаслиги бўлиши зарур ва етарли.

μ оператор (келтириш)ни киритиб оламиз: $\mu: \aleph^\wp \rightarrow \aleph^\wp$, бунда $(\varphi(\eta), \mu(\eta)) = \min_{\vartheta \in \aleph^\wp} (\varphi(\eta), \vartheta)$.

μ оператор аниқ конструктив кўринишга эга бўлиб, агар $\varphi(\eta)$ вектор компоненталарини камайиш тартибида тартибланса, яъни $\varphi^{j_1}(\eta) \leq \varphi^{j_2}(\eta) \leq \dots \leq \varphi^{j_N}(\eta)$, у ҳолда $\mu(\eta)$ векторнинг компоненталари қуйидагича бўлади: $\mu^{j_1}(\eta) = 1, \mu^{j_2}(\eta) = 1, \dots, \mu^{j_i}(\eta) = 1, \mu^{j_{i+1}}(\eta) = 0, \mu^{j_{i+2}}(\eta) = 0, \dots, \mu^{j_N}(\eta) = 0$.

Бошқача қилиб айтганда, $\varphi(\eta)$ векторнинг дастлабки η та минимал компонентасига мос $\mu(\eta)$ векторнинг компоненталари бирга, қолганлари эса нолга тенг бўлади.

$\mu(\eta)$ векторнинг \wp -ноинформатив вектор эканлиги равшан бўлиб, бу вектор учун қуйидаги шарт ўринли.

$$(\varphi(\eta), \mu(\lambda)) = \min \{(\varphi(\eta), \vartheta) \mid \vartheta \in \aleph^\wp\}. \quad (19)$$

2-хосса. Ихтиёрий $\eta (\eta \in \aleph^\wp)$ учун $(\varphi(\eta), \mu(\eta)) \leq 0$ тенгсизлик ўринли.

11-теорема. Агар $I(\eta) = I(\mu(\eta))$ бўлса, у ҳолда $I(\eta) = \min \{I(\vartheta) \mid \vartheta \in \aleph^\wp\}$ бўлади.

11-теорема олинган ечимни оптимал эканлигини таъминлаб беради, яъни $I(\eta)$ функционал \aleph^\wp тўпламининг η элементида ўзининг минимумига эришади.

(5) мезон асосида информатив белгилар мажмуасини аниқлашнинг тартибланиш усули ноинформатив белгилар мажмуасини аниқлашга

қуйидаги оптималлаштириш масалани ечиш учун мослаштириш мумкин.

Фараз қилайлик, a ва b векторлар компоненталарига нисбатан мос равишда қуйидагича тартибланган бўлсин:

$$\frac{a_1}{b_1} \leq \frac{a_2}{b_2} \leq \dots \leq \frac{a_N}{b_N}. \quad (20)$$

У ҳолда $\eta = (\underbrace{1,1,1,\dots,1}_{\varphi}, \underbrace{0,0,\dots,0}_{N-\varphi})$ вектор тартибланиш усули натижасига кўра (18)

масаланинг оптимал ечими ҳисобланади. Бироқ тартибланиш усули ҳар доим ҳам Фишер мезонига нисбатан оптимал ечимни таъминламайди.

Қуйидаги белгиланишларни киритайлик:

$$A = \sum_{i=1}^l a_i, \quad B = \sum_{i=1}^l b_i, \quad \begin{cases} \Delta a_{ij} = a_j - a_i \\ \Delta b_{ij} = b_j - b_i, i = \overline{1, l}, j = \overline{l+1, N} \end{cases}, \quad \eta^0 = \left(\underbrace{1,1,\dots,1}_{\varphi \text{ та}}, \underbrace{0,0,\dots,0}_{N-\varphi \text{ та}} \right).$$

Юқорида келтирилган 1-6-леммаларда $a = \Delta a_{ij}, b = \Delta b_{ij}, c = A, d = B$ деб олсак, у ҳолда $\forall i, j (i = \overline{1, l}, j = \overline{\varphi+1, N})$ лар учун $\begin{cases} A + \Delta a_{ij} \geq 0, \\ B + \Delta b_{ij} > 0 \end{cases}$ бўлиб, леммалардан бири ўринли бўлади.

12-теорема. Тартибланган (20) кетма-кетлик ёрдамида танлаб олинган $\eta^0 = (\underbrace{1,1,\dots,1}_{\varphi \text{ та}}, \underbrace{0,0,\dots,0}_{N-\varphi \text{ та}})$ вектор (18) масаланинг оптимал ечим бўлиши учун 1- ва 3-лемма шартларини қаноатлантирувчи $a = \Delta a_{ij}, b = \Delta b_{ij}$ ларнинг мавжуд бўлмаслиги зарур ва етарли.

Модификацион «Дельта» усули (5) мезон асосида ноинформатив белгилар мажмуасини аниқлашга мослаштирилган бўлиб, қуйидаги теоремага асосланади.

$\forall \eta \in \mathbb{N}^{\varphi}$ танланган бўлсин.

13-теорема. Танланган η вектор (18) масаланинг оптимал ечими бўлиши учун 1-лемма, 3-лемма ва 6-лемма шартларини қаноатлантирувчи $a = \Delta a_{ij}$ ва $b = \Delta b_{ij} (i = \overline{1, l}, j = \overline{l+1, N})$ ларнинг мавжуд бўлмаслиги зарур ва етарли.

η вектор (18) масаланинг ечими бўлмаса, у ҳолда 1-, 3- ва 6-леммалар асосида алмаштиришлар бажарамиз. Алмаштириш жараёни 1-, 3- ва 6-леммалар шартларини қаноатлантирувчи Δa_{ij} ва Δb_{ij} мавжуд бўлмагунча давом эттирилади. Агар Δa_{ij} ва Δb_{ij} 1-, 3- ва 6-леммалар шартларини қаноатлантирувчи Δa_{ij} ва Δb_{ij} мавжуд бўлмаса, 13-теорема натижасига кўра ҳосил қилинган ечим оптимал.

4.2-параграфда 0-тартибли бир жинсли умумий кўринишдаги мезон асосида ноинформатив белгиларни аниқлаш усуллари келтирилган бўлиб, (6) функционалдан ноинформатив белгиларни аниқлашда фойдаланилган.

Дастлаб η векторни $\eta^0 = (\underbrace{1,1,\dots,1}_{\varphi}, \underbrace{0,0,\dots,0}_{N-\varphi})$ кўринишда танлаб,

$$A = \sum_{i=1}^l a_i, \quad B_j = \sum_{i=1}^l b_i^j, \quad j = \overline{1, r} \text{ деб оламиз.}$$

(18) масалани ечиш учун ёрдамчи C вектордан фойдаланамиз. Бу вектор координаталари $c_i = \frac{a_i}{A} - \sum_{j=1}^r \frac{b_j^{(i)}}{B_j}$ формула орқали ҳисобланади.

$\forall \mu \in \mathbb{S}^\rho$ ва η^0 вектор асосида шакллантирилган C векторлар берилган бўлсин.

14-теорема. Танлаб олинган η вектор (18) масаланинг оптимал ечими бўлиши учун $(C, \mu) < 0$ шартни қаноатлантирувчи μ векторнинг мавжуд бўлмаслиги зарур ва етарли.

14-теоремага асосланган усул ноинформатив белгиларни аниқлашнинг «Градиент-1» усули бўлиб, μ векторни шакллантириш жараёни $I(\eta) = I(\mu)$ тенглик бажарилмагунча давом эттирилади.

(8) функционалга оид (18) оптимизация масаласини кўриб чиқайлик. Қуйидаги белгилашларни киритайлик:

$$A = \sum_{i=1}^l a_i, B = \sum_{i=1}^l b_i, C = \sum_{i=1}^l c_i, \begin{cases} \Delta a_{ij} = a_j - a_i, \\ \Delta b_{ij} = b_j - b_i, i = \overline{1, l}, j = \overline{l+1, N}. \end{cases}$$

Фараз қилайлик, $\forall \eta \in \mathbb{S}^\rho$ танланган бўлсин.

15-теорема. Танланган η вектор (18) масаланинг оптимал ечими бўлиши учун 8-лемма шартларини қаноатлантирувчи $a = \Delta a_{ij}, b = \Delta b_{ij}$ ва $c = \Delta c_{ij}$ ($i = \overline{1, l}, j = \overline{l+1, N}$) ларнинг мавжуд бўлмаслиги зарур ва етарли.

(8) мезон асосида ноинформатив белгиларни аниқлашнинг модификацион Дельта усули 15-теоремага асосланади ва у тартиблаш усули каби амалга оширилади.

Бундан ташқари, 4.2-параграфда (8) функционалга мос ноинформатив белгиларни ажратишнинг (18) масалани ечишнинг “Градиент-2” усулини кўриб чиқилган.

Ёрдамчи $D(d_1, d_2, \dots, d_n)$ векторни киритиб оламиз, унинг компоненталари қуйидагича аниқланади: $d_i = \frac{a_i}{A} - \frac{b_i}{B} - \frac{c_i}{C}$. Бу ерда $A = (a, \eta), B = (b, \eta), D = (d, \eta), A_1 = (a, \mu), B_1 = (b, \mu), D_1 = (d, \mu)$.

У ҳолда (8) функционалнинг мос қийматлари $I(\eta) = \frac{A}{BD}$ ва $I(\mu) = \frac{A_1}{B_1 D_1}$ га тенг бўлади.

Танлаб олинган λ вектор ва у асосида шакллантирилган C вектор учун қуйидаги теорема ўринли.

16-теорема. $I(\eta) < I(\mu)$ бўлиши учун $(C, \mu) > 0$ бўлиши етарли.

17-теорема. Агар $I(\eta) = I(\mu)$ бўлса, у ҳолда $I(\eta) = \min \{I(\vartheta) | \vartheta \in \mathbb{S}^\rho\}$.

17-теорема олинган ечимнинг \mathbb{S}^ρ соҳада оптималлиги таъминлайди.

16- ва 17-теоремалар ёрдамида (11) функционални оптималлаштирувчи итерацион усул асосланади. Бу усулнинг биринчи қадамида \wp -ноинформатив η вектор ихтиёрий танланади. Қулай бўлиши учун η векторни $\eta = \overbrace{(1, 1, \dots, 1)}^\wp, 0, 0, \dots, 0$ кўринишда олинади. Кейинги ҳар бир қадамда янги η вектор олдингисига μ векторни қуриш асосида шакллантирилади ва

итерацион жараён $(C, \eta) > 0$ бўлса тўхтатилади ва η -оптимал ечим.

k -тартибли, аралаш бир жинсли ёки бир жинсли бўлмаган мезонлар 0-тартибли ва k -тартибли мезонлар комбинациялари шаклида ифодаланади ва улар асосида ноинформатив белгилар фазосини шакллантириш усули ҳам ноинформатив белгиларни аниқлашнинг “Градиент-2” усули каби амалга оширилади. Бунда йўналтирувчи вектор мезонларга мос йўналтирувчи векторлар комбинациялари асосида шакллантирилади.

Диссертациянинг “**Тажрибавий тадқиқотлар ва амалий тадбиқ**” деб номланган бешинчи боби диссертацияда таклиф этилган информатив белгиларни танлаш усул ва алгоритмларининг тажрибавий тадқиқини ўтказиш учун фойдаланиладиган ўқув ва назорат танланмалари ҳамда белгилар фазосини шакллантириш услубиятини акс эттиради. Бундан ташқари, мазкур бобда “MDIB” дастурлар мажмуасининг тавсифи ва уни амалий тадбиқи келтирилган.

5.1-параграфда қуйидаги талабларга жавоб берувчи ўқув ва назорат танланмаларидан иборат бошланғич маълумотлар базасини шакллантирилди:

- ўқув ва назорат танланмалари бир хил қонун ва қоидалар асосида яратилган бўлсин;
- ўқув ва назорат танланмалар объектларининг синфлари сони камида учта бўлсин;
- информатив белгилар сони ℓ ($1 \leq \ell \leq N$)та ва бу белгилар олдиндан маълум бўлсин.

5.2-параграф N ўлчовли R^N белгилар фазосида қурилган учта K_1, K_2 ва K_3 синфлар учун тажрибавий тадқиқотлар натижаларини акс эттиради. N ўлчов олдиндан берилган бирор бир унча катта бўлмаган ℓ сонига қаррали қилиб олинган, яъни $N = \ell * h$ ($h \in N$). Бошланғич $x = (x_1, x_2, \dots, x_N)$ белгилар мажмуасини ҳар бир гуруҳда ℓ тадан белгини мужассамлаштирган белгилар гуруҳига ажратилган, яъни $(x_1, \dots, x_\ell, x_{\ell+1}, \dots, x_{2\ell}, x_{2\ell+1}, \dots, x_{(h-1)\ell+1}, \dots, x_{h\ell})$. Бу белгилар гуруҳи шундай ажратилганки, белгилар гуруҳлар бўйича ўзаро боғлиқсиз бўлиб, гуруҳлар ичида кучли боғланган, яъни бирор бир гуруҳ ичидаги ихтиёрий белгини чиқариб юборилса, у ҳолда ўша гуруҳ белгилари орасидаги боғлиқлик, яъни белгилар мажмуасининг информативлиги бирданига ёмонлашиб кетади.

Тадқиқотлар (8) функционал учун оптимизация масаласини “Дельта-2” ва тўла танлов усуллари асосида ечиш билан боғлиқ. “Дельта-2” усулида олинган информатив белгилар мажмуасини аниқлаш натижалари тўла танлов усулида олинган натижалар билан тўла мос тушди. Бунда тўла танлов усулида итерациялар сони C_N^ℓ ни, “Дельта-2” усули учун ℓ га тенг бўлди.

Диссертациянинг 5.3-параграфи мавжуд ва диссертацияда таклиф этилган маълумотларга дастлабки ишлов бериш алгоритмлари асосида яратилган “MDIB” дастурий таъминотни тавсифига бағишланади.

Диссертациянинг 5.4-параграфи дастурий таъминотнинг амалий масалаларда қўлланилишига бағишланган бўлиб, унда КНК-270

комплексдан олинган маълумотларга дастлабки ишлов бериш ҳамда асосий кўрсаткичлари орқали унинг ишлаш хавфсизлигини таъминлаш масаласи кўриб чиқилган.

Тадқиқот объекти сифатида Навоий тоғ-металлургия комбинатида ўрнатилган КНК-270 конвейерининг қуйидаги уч: хавфли, ўртача, хавфсиз ҳолатлари тадқиқ қилинди. Комплексининг хавфли ҳолатини характерловчи сигналлар 1-синфни K_1 , ўртача ҳолатини 2-синфни K_2 , хавфли ҳолатини 3-синфни эса K_3 ни ифодалайди.

Ҳар бир синф объектларни характерловчи белгилар сони 9 тадан иборатдир.

Ҳар бир синфда объектлар сони бир хил бўлиб, 5056тани ташкил этади. Шундай қилиб, ҳар бир объектни $x_{ij} = (x_{ij}^1, x_{ij}^2, \dots, x_{ij}^9)$ вектор кўринишида ифодалаш мумкин, бу ерда x_{ij}^k - i - синф j -объектнинг k -белгиси; $k = \overline{1,9}$; $j = \overline{1,5056}$; $i = \overline{1,3}$.

КНК-270 комплекси асосий кўрсаткичларини аниқлаш масаласининг математик қўйилиши қуйидагича:

$$\begin{cases} I(\lambda) = \frac{(a, \lambda)}{(b, \lambda)(c, \lambda)(d, \lambda)} \rightarrow opt, \\ \lambda \in \Lambda^\ell. \end{cases}$$

Информатив белгилар мажмуасини аниқлагандан сўнг ($\ell = \overline{1,9}$ учун), “ k -яқин қўшнилар” усули ёрдамида тимсолларни аниқлаш масаласи ҳал этилди, бу ушбу белгилар тўпламларининг ҳар бирининг “фойдалилик” даражасини ўқув танланма объектларини таниб олиш сифатига таъсири нуқтаи назаридан баҳолаш имконини берди.

Ушбу масалани ечиш натижасида нисбатан информатив бўлган белгилар мажмуаси аниқланди x_2, x_5, x_8 , бу ерда x_2, x_5 ва x_8 - уч компонентали $D_i, i = \overline{1,3}$ датчикларнинг вертикал компоненталаридан келадиган сигнал қийматларини ифодалайди.

Олинган натижалар асосида соҳа мутахассислари томонидан КНК-270 комплексининг кўчириш жараёнида юзага келадиган аварияли ҳолатларни олдини олиш бўйича тавсиялар ишлаб чиқилди.

ХУЛОСА

“Маълумотларга дастлабки ишлов бериш усул ва алгоритмлари” мавзусидаги диссертация иши бўйича олиб борилган тадқиқотлар натижасида қуйидаги хулосалар тақдим этилди:

1. Маълумотлар таҳлили масалаларида белгилар фазоси ўлчамини камайтириш муаммоси формаллаштирилди. Бу информатив белгилар аниқлаш масаласини шакллантиришга хизмат қилади. Белгилар информативлигининг ва ноинформативлигининг эвристик мезонлари ишлаб чиқилди. Ушбу мезонлар объектни характерловчи асосий кўрсаткичларни аниқлаш имконини беради.

2. Ягона информативлик мезонидан фойдаланиб, информатив ва ноинформатив белгилар мажмуасини аниқлаш усул ва алгоритмлари яратилган. Мазкур усул ва алгоритмлар информатив белгилар ишончилигини оширишга хизмат қилади.

3. Дастлабки белгилар фазоси ўлчамини камайтириш масаласи экстремал қийматларини берадиган шундай l -ўлчовли ($l < N$) белгилар қисм системаси топишдан иборат оптималлаштириш масаласига келтирилди. Мазкур оптималлаштириш масаласининг ечими кам сонли белгиларда объектларни информатив таснифлашга хизмат қилади.

4. Фишер ва Горелик эвристик информативлик мезонлари хоссалари таҳлили асосида бу икки мезоннинг принципиал фарқи аниқланди. Бунда Горелик функционали хусусий мезонларини жамлашни амалга оширмайди, фақат битта максимал ёки минимал бўлганини танлаш имконини беради. Бу эса синфлар ажралиши ишончилигини ошишига хизмат қилади.

5. Фишер информативлик мезонининг хусусиятлари аниқланди ҳамда кейинчалик ушбу хусусиятлардан информатив белгилар мажмуасини аниқлаш усуллари ишлаб чиқиш учун тузилмавий асос сифатида фойдаланилди. Хусусан, берилган l -информатив белгилар қисм системаси учун Фишер мезонининг қуйи ва юқори чегаралари аналитик аниқланди. Синфлар орасидаги масофани характерловчи вектор мос компоненталари объектларнинг синфлар орасидаги тарқоклигини характерловчи вектор компоненталари муносабатининг максимал қийматини билдирувчи белги ҳар доим информативлик мезонининг максимал қийматини таъминловчи l -информатив белгилар қисм системасида мавжуд бўлади. Фишер мезонининг ушбу хоссаси N -ўлчовли белгилар фазосидан $N-l$ -ўлчовли фазога ўтиш имкониятини беради.

6. Фишер информативлик мезони учун информатив белгилар мажмуасини аниқлашнинг регуляр, рекуррент, қисмий танлов, усуллари тақлиф этилди. Информатив белгилар мажмуасини аниқлашнинг Фишер информативлик мезонига оид оптималлаштириш масаласини регуляр усулда ечишда $l(\lambda)$ функционални тезкор ўсиш йўналишини кўрсатувчи вектор функциядан фойдаланиши кўрсатиб берилди. Тартибланиш бажарилганда регуляр усулдан фарқли бўлган, чекли қадамларда $((2N-l+1)/2)$ оптимал ечимни таъминловчи рекуррент усули ишлаб чиқилди. Бу оптималлаштириш

масалани қисмий танлов усулини орқали ечиш имконини беради.

7. Тартиблаш усули ҳар доим ҳам Фишер информативлик мезонига оид оптималлаштириш масаласининг оптимал ечимини таъминламаслиги кўрсатиб берилди ва оптималлик шартлари аниқланди. Бу шартлар субоптимал ечимлардан оптимал ечимни олиш имконини беради.

8. Танланган ечимнинг оптималлик зарурий ва етарлилик шартлари аниқланди ва бу шартлар асосида тезкор “Дельта” усули ишлаб чиқилди. “Дельта” усули Фишер типигаги содда мезон асосида информатив белгилар фазосини шакллантиришдан олинган натижаларни оптималлигини текшириш ва бошқа усуллардан олинган субоптимал ечимдан фойдаланиб тезкорлик билан оптимал ечим олишга хизмат қилади.

9. Бир жинсли нолинчи тартибли умумлашган информативлик мезонлари асосида информатив ҳамда ноинформатив белгилар мажмуасини аниқлаш усул ва алгоритмлари ишлаб чиқилган. Ушбу усул ва алгоритмдан олинган натижалар объектларни информатив таснифлашга хизмат қилади.

10. Бир жинсли k -тартибли умумлашган информативлик мезонлари асосида информатив ҳамда ноинформатив белгилар мажмуасини аниқлаш усуллари ва алгоритмлари ишлаб чиқилган. Мазкур усул ва алгоритмлар ҳар бир синф орасидаги объектлар жойлашуви қонуниятларини ҳисобга олган ҳолда объектларни информатив таснифлаш имконини беради.

11. Аралаш типли информативлик мезонлари асосида информатив ва ноинформатив белгилар мажмуасини аниқлаш усуллари ва алгоритмлари ишлаб чиқилган. Мазкур усул ва алгоритмлар ўлчамлари турлича бўлган белгилар фазосидаги объектларни информатив таснифлаш имконини беради.

12. Эвристик информативлик мезони минимизацияси билан боғлиқ ноинформатив белгилар мажмуасини аниқлаш усул ва алгоритмлари ишлаб чиқилди. Ушбу усулларнинг қуриш тамойиллари мазмунига кўра информатив белгилар мажмуасини аниқлаш усуллари қуриш тамойилларига яқин бўлиб, уларнинг фарқи информативлик мезони максимизация масаласи (информатив белгилар мажмуасини аниқлашда) ўрнига минимизация масаласи (ноинформатив белгилар мажмуасини аниқлашда) ечилади. Ноинформатив белгилар мажмуасини аниқлаш усул ва алгоритмлари информатив белгилар мажмуасини аниқлаш имконини беради.

**РАЗОВЫЙ НАУЧНЫЙ СОВЕТ ПРИ НАУЧНОМ СОВЕТЕ
DSc.27.06.2017.Т.07.01 ПО ПРИСУЖДЕНИЮ УЧЕНЫХ СТЕПЕНЕЙ
ПРИ ТАШКЕНТСКОМ УНИВЕРСИТЕТЕ ИНФОРМАЦИОННЫХ
ТЕХНОЛОГИЙ**

**НАУЧНО-ИННОВАЦИОННЫЙ ЦЕНТР ИНФОРМАЦИОННО-
КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ ПРИ ТАШКЕНТСКОМ
УНИВЕРСИТЕТЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ
ИМЕНИ МУХАММАДА АЛ-ХОРАЗМИЙ**

МАМАТОВ НАРЗИЛЛО СОЛИДЖОНОВИЧ

**МЕТОДЫ И АЛГОРИТМЫ ПРЕДВАРИТЕЛЬНОЙ ОБРАБОТКИ
ДАНЫХ**

05.01.03 - Теоретические основы информатики

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

Ташкент – 2018

Тема докторской диссертации по техническим наукам (DSc) зарегистрирована в Высшей аттестационной комиссии при Кабинете Министров Республики Узбекистан за №B2017.2.DSc/T86.

Диссертация выполнена в Научно-инновационном центре информационно-коммуникационных технологий при Ташкентском университете информационных технологий.

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

Научный консультант: **Фазылов Шавкат Хайруллаевич**
доктор технических наук, профессор

Официальные оппоненты: **Кобулов Анвар Васильевич**
доктор технических наук, профессор

Рустамов Насим Тулагенович
доктор технических наук, профессор (Казахстан)

Нурмухамедов Толаниддин Рамзидинович
доктор технических наук

Ведущая организация: **Ташкентский государственный технический университет**

Защита диссертации состоится «__» _____ 2018 г. в ____ часов на заседании Научного совета DSc.27.06.2017.T.07.01 при Ташкентском университете информационных технологий. (Адрес: 100202, 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 №__). (Адрес: 100202, г.Ташкент, ул. Амира Темура, 108. Тел.: (99871)238-65-44).

Автореферат диссертации разослан «__» _____ 2018 года.
(протокол рассылки №__ от «__» _____ 2018 г.).

Р.Х.Хамдамов

Председатель научного совета по присуждению
учёных степеней, д.т.н., проф.

Ф.М.Нуралиев

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

М.М.Камилов

Председатель научного семинара при научном совете по
присуждению ученых степеней, д.т.н., академик

ВВЕДЕНИЕ (аннотация диссертации доктора наук (DSc))

Актуальность и востребованность темы диссертации. В мире огромное внимание уделяется разработке методов компьютерного анализа информации, классификационной обработке данных и распознаванию образов. В этой области разработка систем распознавания и классификации, технических и промышленных приложений технической диагностики, робототехнических систем, материаловедений, управления технологическими процессами, промышленными объектами является одной из важных задач. К настоящему времени разработаны и исследованы разнообразные методы распознавания и классификации, такие как, детерминированные, вероятностные, логические, структурные. В зарубежных странах, в том числе в США, Германии, России, Китае, Японии, Южной Корее и др., большое внимание уделяется решению теоретических и практических задач распознавания образов и классификационной обработки данных.

В мире ведутся исследовательские работы по созданию систем искусственного интеллекта, ориентированных на решение задач управления объектов, процессов и явлений, прогнозирования, идентификации и основанных на использовании алгоритмов распознавания образов. В этой связи, в том числе, важной задачей является усовершенствование методов и алгоритмов устранения пропусков, шумов и выбросов, унификация типов переменных, разработка методов предварительной обработки исходных данных, создание методов и алгоритмов информативного описания объектов, распознавания, выбора информативных признаков, определение наборов неинформативных признаков.

В Республике информационно-коммуникационные технологии широко применяются в корпоративное управление в экономике, социальной сфере, народном хозяйстве и других областях. В этом направлении особое внимание уделяется предварительной обработке и компьютерному анализу данных, в том числе, разработке методов и алгоритмов формирования информативного описания объектов, явлений и процессов. В Стратегии действий по дальнейшему развитию Республики Узбекистан в 2017-2021 гг. отмечены задачи, в том числе, по «... внедрению современных методов и стандартов в корпоративное управление, ... внедрению информационно-коммуникационных технологий в экономику, социальную сферу, систему управления»¹. Для реализации подобных задач важным является создание систем поддержки принятия решений с использованием методов распознавания образов. В этой связи, важное значение имеет совершенствование методов предварительной обработки исходных данных, используемых для принятия решений.

¹ Указ Президента Республики Узбекистан от 7 февраля 2017 г. № УП-4947 «О Стратегии действий по дальнейшему развитию Республики Узбекистан».

Данное диссертационное исследование в определенной степени служит для выполнения задач, предусмотренных указами Президента Республики Узбекистан №УП-4947 от 7 февраля 2017 года “О стратегии действий по дальнейшему развитию Республики Узбекистан”, №УП-5264 от 29 ноября 2017 года “Об образовании Министерства инновационного развития Республики Узбекистан”, постановлениями Президента №ПП-1730 от 21 марта 2012 года “О мерах по дальнейшему внедрению и развитию современных информационно-коммуникационных технологий”, №ПП-2158 от 3 апреля 2014 года “О мерах по дальнейшему внедрению информационно-коммуникационных технологий в реальном секторе экономики”, №ПП-2898 от 18 апреля 2017 года “О мерах по коренному совершенствованию деятельности органов внутренних дел в сфере расследования преступлений”, 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 IV. “Развитие информатизации и информационно-коммуникационных технологий”.

Обзор зарубежных научных исследований по теме диссертации². Научные исследования направленные на разработку методов и систем распознавания образов, а также формирования информативного описания исследуемых объектов, осуществляются в ведущих научных центрах мира и высших образовательных учреждениях, в том числе, в Carnegie Mellon University, University of California, Department of Electrical and Computer Engineering University of Texas (США), Darmstadt University of Applied Sciences (Германия), Attar Software Ltd (Великобритания), SIPINA, University of Lyon (Франция), Izmir University (Турция), Institute of Cybernetics of National Academy of Sciences of Ukraine (Украина), United Institute of Informatics Problems, Belarusian State University of Informatics and Radioelectronics (Белоруссия), Azerbaijan Association of Zadeh’s Legacy, Department of EECS, Department of Business Administration (Азербайджан), Вычислительном центре Российской академии наук, Институте проблем передачи информации, Санкт-Петербургском институте информатики и автоматизации, Московском Государственном Университете, Институте прикладной математики, Томском Государственном Университете (Российская Федерация), Институте автоматизации (Китай), Tokyo Institute of Technology (Япония), Национальном Университете Узбекистана (Узбекистан).

В результате исследований, проведенных в мире по созданию методов

² Обзор научных исследований по теме диссертации составлен на основании <https://elibrary.ru/item.asp?id=21681096>, <https://studfiles.net/1376730/>, <https://cyberleninka.ru/.../obzor-metodov-raspoznavaniya-simvolov>, <http://intellect-tver.ru/?p=165>, <http://www.nickart.spb.ru/analysis/market.php>, <http://www.cedar.buffalo.edu/~srihari/CSE555/>, <https://www.ibm.com/developerworks/ru/library/ba-data-mining-techniques/index.html>, www.open.edu/openlearn-create/mod/resource/view.php?id, <http://ieeexplore.ieee.org/document/824819/> и других источников.

и алгоритмов распознавания образов и предварительной обработке данных, а также совершенствование систем распознавания и классификации получен ряд результатов, в том числе созданы методы и алгоритмы устранения шумов в данных (Международная академия информатизации, Российская Федерация; Stanford University, США; Transport and Telecommunication Institute, Латвия); разработаны методы и алгоритмы сглаживания и фильтрации данных (Massachusetts Institute of Technology, International Institute of Forecasters, США; Центральный экономико-математический институт, Российская Федерация); разработаны методы и алгоритмы унификации типов признаков, характеризующих объектов (Московский Государственный Технический Университет им.Н.Э.Баумана, Санкт-Петербургский институт информатики и автоматизации, Центральный экономико-математический институт, Российская Федерация); разработаны методы и алгоритмы снижения размерности признакового пространства (Институт математики СО РАН, Институт проблем управления, Российская Федерация; Bell Labs, США; Corporation NEC, Япония); создано программное обеспечение распознавания речи (Nuance Communications, Google, США; Yandex, Российская Федерация); созданы системы распознавания личности (Kairos, Англия; NVIDIA, Microsoft, США); созданы системы распознавания текста (АВВУУ, Российская Федерация).

В мире по дальнейшему развитию существующих и созданию новых методов анализа данных и распознавания образов осуществляется по следующим перспективным направлениям: разработка методов и алгоритмов, обеспечивающих интеллектуализацию систем распознавания, функционирующих при весьма ограниченных объемах исходной информации; разработка методов и алгоритмов распознавания при большой размерности признакового пространства; разработка методов и алгоритмов определения объектов в изображениях; создание современных систем распознавания речи, текста, идентификация личности с использованием технологий «мягких» вычислений, обеспечивающих интегрированную обработку входной информации.

Степень изученности проблемы. Научные основы предварительной обработки данных, в том числе, снижения размерности признакового пространства в задачах анализа данных, в частности, в распознавании образов, разработаны и изложены в работах С.А.Айвазяна, Б.Н.Адамовского, В.В.Александрова, С.В.Абламейко, В.И.Васильева, Н.Д.Горского, Н.Г.Загоруйко, Г.С.Лбова, Н.А.Игнатьева, Ш.Ю.Раудиса, В.В.Старовойтова, I.Marill, D.M.Green и др. В работах Ю.И.Журавлева, М.М.Камилова, З.Т.Адыловой, Р.А.Лутфуллаева, Ш.Е.Туляганова исследованы свойства использования меры важности исходного признака, представляющей собой обработанную надлежащим образом степень уменьшения так называемых «голосов» при удалении этого признака. Авторами метода эта мера названа информационным весом признака.

Работы А.Л.Горелика, Г.И.Кутина, К.А.Чепониса, Д.А.Жвинерайте, Б.С.Бусыгина, Л.В.Мирошниченко, R.Fisher посвящены исследованию

эвристических критериев информативности признаков, а также разработке методов перебора, направленных на формирование информативного описания объектов на основе этих критериев. В дальнейшем это направление было развито в работах М.М.Камилова, Ш.Х.Фазылова, А.Х.Нишанова, в которых были предложены поисковые методы определения информативных наборов признаков и введено понятие ℓ -информативного вектора признаков. В работах Р.Хамдамова впервые предложен аналитический метод определения информативных признаков на основе эвристических критериев информативности Фишера.

В распознавании образов почти во всех исследованиях, связанных с решением задачи снижения размерности признакового пространства, требуется выбор наиболее подходящей для конкретной задачи комбинации метода и критерия, т.е. фиксированием для этой задачи такой пары метода и критерия, при использовании которых минимизируются вычислительные затраты и улучшается качество распознавания. Поэтому для распространенных критериев информативности разработка математически обоснованных соответствующих методов является актуальной проблемой, которая в настоящее время в достаточной степени не решена.

Связь диссертационного исследования с планами научно-исследовательских работ научно-исследовательского учреждения, где выполнена диссертация. Диссертационное исследование выполнено в рамках проектов плана научно-исследовательских работ Научно-инновационного центра информационно-коммуникационных технологий по следующим темам: ФА-А17-Ф010: «Разработка алгоритмического и программного обеспечения многоуровневой системы биометрической идентификации личности» (2009-2011); А5-ФА-Ф022: «Разработка и практическое применение гибридных алгоритмов и программного обеспечения биометрических решений» (2012-2014); А5-003: «Разработка и практическое применение алгоритмического и программного обеспечения системы анализа рукописных текстов» (2015-2017); БВ-М-Ф4-003: «Методы синтеза распознающих операторов, основанные на оценке взаимосвязанности признаков» (2017-2020).

Целью исследования является разработка методов и алгоритмов формирования информативных описаний объектов с использованием эвристических критериев информативности как частного, так и обобщенного видов.

Задачи исследования:

разработка эвристических критериев информативности и неинформативности признаков и оценка их эффективности;

определение свойств критериев информативности Фишера, используемых при создании методов определения информативных и неинформативных наборов признаков;

разработка методов и алгоритмов определения информативных и неинформативных наборов признаков на основе единого критерия информативности;

разработка методов и алгоритмов определения информативных и неинформативных наборов признаков с использованием обобщенных однородных критериев информативности нулевого и k -го порядка;

разработка методов и алгоритмов определения информативных и неинформативных наборов признаков с использованием критериев информативности смешанного типа;

создание правила разбиения исходной совокупности признаков на группы информативных, малоинформативных и неинформативных признаков с использованием эвристических критериев;

создание алгоритмического и программного обеспечения предложенных методов определения информативных и неинформативных наборов признаков.

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

Предметом исследования являются методы, алгоритмы и эвристические критерии формирования информативного описания объектов.

Методы исследования. Теоретические исследования в работе базировались на методах математического анализа, дискретной математики и распознавания образов, в основе экспериментов использовалось компьютерное моделирование процесса формирования пространства информативных признаков с последующим распознаванием заданных объектов в этом пространстве.

Научная новизна диссертационного исследования заключается в следующем:

разработаны эвристические критерии информативности и неинформативности признаков и оценена их эффективность;

определены свойства критериев информативности Фишера, используемые при создании методов определения информативных и неинформативных наборов признаков;

усовершенствованы методы и разработаны алгоритмы выбора информативных и неинформативных наборов признаков с использованием единого критерия информативности на основе гипотезы компактности;

разработаны методы и алгоритмы выбора информативных и неинформативных наборов признаков с использованием обобщенных однородных критериев информативности нулевого и k -го порядка;

усовершенствованы методы и разработаны алгоритмы выбора информативных и неинформативных наборов признаков с использованием критериев информативности смешанного типа;

создано правило разбиения исходной совокупности признаков на группы информативных, малоинформативных и неинформативных признаков с использованием эвристических критериев.

Практические результаты исследования заключаются в следующем:

разработаны поисковые методы определения информативных и неинформативных наборов признаков, результаты реализации которых совпадают с результатами метода полного перебора при существенно меньших вычислительных затратах;

для каждого вида эвристического критерия информативности разработан соответствующий метод определения информативных и неинформативных наборов признаков, что существенно упрощает выбор для конкретной практической задачи наиболее подходящей пары «критерий-метод».

Достоверность результатов исследования. Достоверность результатов исследования обосновывается корректностью постановки задачи снижения размерности признакового пространства, как задачи дискретной оптимизации, строгостью математических выкладок, представленных в виде доказанных теорем и лемм, 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 распознавания образов в части разработки математически строгих методов снижения размерности признакового пространства с использованием эвристических критериев информативности признаков при формировании информативного описания объектов, явлений и процессов. Разработанные методы и алгоритмы обосновываются их широким применением при разработке систем распознавания и классификации, используемых в технической диагностике, робототехнических системах, материаловедении, управлении технологическими процессами и производствами.

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

Внедрение результатов исследования. На основе программного обеспечения, созданного на базе существующих, а также предложенных в работе методов и алгоритмов предварительной обработки данных:

разработанный программный комплекс, позволяющий осуществлять предварительную обработку и хранение данных, получаемых с вибродатчиков внедрен в Центральное рудоуправление Навоинского горно-металлургического комбината (справка Министерства по развитию информационных технологий и коммуникаций №33-8/8663 от 21 декабря 2017 г.). На основе использования результатов научного исследования сократились простои конвейерного комплекса КНК-270 на 10% за счет оперативного оповещения об опасных критических ситуациях, возникающих в процессе работы данного комплекса;

разработанные алгоритмы повышения качества речевых сигналов на основе их предварительной обработки и фильтрации внедрены в Следственном департаменте Министерства внутренних дел Республики Узбекистан (справка Министерства по развитию информационных

технологий и коммуникаций №33-8/8663 от 21 декабря 2017 г.). На основе использования результатов научного исследования время обработки речевых сигналов при их переводе в текстовую информацию сократился на 20%;

разработанные программные средства интеллектуального анализа данных в системе электронного документооборота внедрены в Министерстве высшего и среднего специального образования в рамках системы “Электронное министерство” (справка Министерства по развитию информационных технологий и коммуникаций №33-8/8663 от 21 декабря 2017 г.). На основе использования результатов научного исследования эффективность процесса обработки данных о профессорско-преподавательском составе высших учебных заведений и формирования соответствующих предложений через модуль кадры системы “Электронное министерство” повысилась в 2 раза.

разработанные программные средства формирования и предварительной обработки данных с целью определения основных показателей производимой продукции внедрены в ООО “INTERVETFARM” (справка Министерства по развитию информационных технологий и коммуникаций №33-8/8663 от 21 декабря 2017 г.). На основе использования результатов научного исследования обеспечилась устойчивость качества производимой продукции за счет определения показателей сырья, которые необходимо постоянно контролировать, что дало возможность сократить на 20% время, требуемое для контроля качества.

программные средства предварительной обработки, сглаживания и фильтрации данных, обеспечивающие точность контурных линий за счет устранения различных выбросов внедрены в ООО “Inter sharm plast” (справка Министерства по развитию информационных технологий и коммуникаций № 33-8/8663 от 21 декабря 2017 г.). На основе использования результатов научного исследования время производства продукции на основе использования специально кодированного файла на станке, отражающего основные (информативные) точки, выделенные на контурных линиях сократился на 15%, что также привело к экономии сырья среднем на 20%. эффективность производства увеличилась в среднем в 1,2 раза.

Апробация результатов исследования. Теоретические и практические результаты диссертационного исследования докладывались и обсуждались на 7 международных и 20 республиканских научно-практических конференциях.

Опубликованность результатов исследования. По теме диссертации опубликовано 43 научных работ. Из них 1 монография, 14 научных статей, в том числе 11 в республиканских и 3 в зарубежных журналах рекомендованных Высшей аттестационной комиссией Республики Узбекистан для публикации основных научных результатов докторских диссертаций, 27 докладов и тезисов, также получены 2 свидетельства об официальном регистрации программы для ЭВМ.

Структура и объем диссертации. Диссертация состоит из введения, пяти глав, заключения, списка использованной литературы и приложений. Объем диссертации составляет 175 страниц.

ОСНОВНОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ

Во введении обоснована актуальность темы диссертации в соответствии с приоритетными направлениями развития науки и технологий Республики Узбекистан. Сформулированы цель и задачи, указаны объект и предмет исследования, изложены научная новизна и практические результаты исследования. Обоснована достоверность полученных результатов, раскрыта теоретическая и практическая значимость полученных результатов. Приведены перечень внедрений результатов исследования в практику, сведения об опубликованных работах и структура диссертации.

В первой главе диссертации **“Проблемы предварительной обработки данных”** проведен анализ задач повышения качества исходных данных за счет устранения пропусков, шумов и выбросов, а также унификации типов переменных путем сведения различных типов признаков к признакам, измеренным в одготипных шкалах.

Низкое качество данных определяется, в основном, пропущенными значениями признаков, а также выбросами и шумами.

Ситуация с пропущенными данными значительно осложняет применение математических методов, так как смещение основных статистических характеристик, например, математического ожидания или дисперсии, возрастает прямо пропорционально числу пропусков.

На сегодняшний день существует множество путей решения проблемы неточных данных. В диссертационной работе подробно рассмотрены наиболее распространенные методы, в частности, методы исключения некомплектных объектов из исходной выборки, заполнения пропусков средними по столбцу значениями, сплайн-интерполяции, а также алгоритм ZET, RESAMPLING метод и другие.

В этой же главе рассмотрены наиболее распространенные методы сглаживания и фильтрации, используемые для решения проблемы анализа данных с выбросами и шумами.

Если результаты экспериментальных наблюдений представлены значениями признаков различного типа, то такие данные называются разнотипными.

Отсутствие возможности совместного анализа разнотипных показателей приводит к тому, что некоторые классы экспериментальных данных вообще не могут быть обработаны, результаты же, полученные после отдельного анализа признаков различных типов, не отражают всего многообразия внутренних связей исследуемого явления и поэтому не представляют практического интереса.

Преодоление указанных трудностей стало возможным благодаря разработке методов обработки разнотипных данных, базирующихся на идее оцифровки неколичественных переменных. Суть данной идеи заключается в приписывании градациям качественных и номинальных признаков действительных числовых значений (числовых меток). Основным преимуществом такого подхода является то, что он позволяет распространять на разнотипные данные многие математические методы анализа числовой информации.

$$f(x^i) = \begin{cases} 0, & \text{если признак } x^i \text{ исключается;} \\ 1, & \text{если признак } x^i \text{ оставляется.} \end{cases}$$

В этом случае новая система признаков формируется как подмножество множества исходных признаков.

Методы, основанные на использовании преобразований типов (2) и (3), называются методами снижения размерности исходного пространства признаков.

Следует отметить, что методы и алгоритмы снижения размерности признакового пространства, предложенные в диссертации, базируются на преобразовании (3).

Первая глава диссертации завершается анализом эвристических критериев и методов, используемых при решении задач снижения размерности признакового пространства. На основе результатов этого анализа сформулированы цель и задачи диссертационного исследования.

Во второй главе диссертации “**Эвристические критерии информативности признаков и их эффективность**” изложены частные и обобщенные виды эвристических критериев информативности признаков. Кроме того, на основе исследования критерия информативности Фишера определены его свойства и эффективность.

В параграфе 2.1 приведены основные понятия, обозначения и необходимые определения.

Допустим, обучающая выборка задана объектами $x_{11}, x_{12}, \dots, x_{1m_1}, x_{21}, x_{22}, \dots, x_{2m_2}, \dots, x_{r1}, x_{r2}, \dots, x_{rm_r}$, для которых известно, что каждая группа объектов $x_{p1}, x_{p2}, \dots, x_{pm_p}$ принадлежит к определенному классу $X_p, p = \overline{1, r}$. Каждый объект x_{pi} является N -мерным вектором числовых признаков, т.е. $x_{pi} = (x_{pi}^1, x_{pi}^2, \dots, x_{pi}^N)$.

Для заданной обучающей выборки объектов $x_{p1}, x_{p2}, \dots, x_{pm_p} \in X_p, p = \overline{1, r}$, где x_{pi} -вектор в N -мерном признаковом пространстве, введем вектор $\lambda = (\lambda^1, \lambda^2, \dots, \lambda^N)$, $\lambda^k \in \{0, 1\}, k = \overline{1, N}$, который, как было отмечено выше, однозначно характеризует определенную подсистему признаков. Компоненты вектора λ , равные единице, указывают на наличие соответствующих признаков в данной подсистеме, а нулевые компоненты свидетельствуют об отсутствии соответствующих признаков.

Определение 1. Усечением пространства $R^N = \{x = (x^1, x^2, \dots, x^N)\}$ по λ назовем пространство $R^N|_{\lambda} = \{x|_{\lambda} = (\lambda^1 x^1, \lambda^2 x^2, \dots, \lambda^N x^N)\}$.

Под усеченными расстояниями между двумя объектами $x, y \in R^N$ будем понимать евклидово расстояние между $x|_{\lambda}, y|_{\lambda}$ в $R^N|_{\lambda}$ т.е.

$$\|x - y\|_{\lambda} = \sqrt{\sum_{k=1}^N \lambda^k (x^k - y^k)^2}.$$

Определение 2. Назовем вектор λ ℓ -информативным, если сумма его компонентов равна ℓ , т.е. $\sum_{i=1}^N \lambda^i = \ell$.

Для каждой подсистемы признаков, заданной ℓ -информативным вектором λ , определено свое ℓ -мерное признаковое подпространство. В каждом из этих подпространств введем евклидову норму относительно усечения по λ $\|x\|_\lambda = \sqrt{\sum_{j=1}^N \lambda^j (x^j)^2}$. Обозначим $\bar{x}_p = \frac{1}{m_p} \sum_{i=1}^{m_p} x_{pi}$, $p = \overline{1, r}$, где \bar{x}_p - усредненный объект класса X_p .

Введем функцию

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

Функция $S_p(\lambda)$ характеризует средний разброс объектов класса X_p в подпространстве признаков, заданных вектором λ . Зададим критерий информативности подсистем признаков в виде функционала

$$I_1(\lambda) = \frac{\sum_{p,q=1}^r \|\bar{x}_p - \bar{x}_q\|_\lambda^2}{\sum_{p=1}^r S_p^2(\lambda)}. \quad (4)$$

Функционал (4) является некоторым обобщением функционала Фишера.

Введем обозначения $a = (a^1, a^2, \dots, a^N)$ и $b = (b^1, b^2, \dots, b^N)$, здесь $a^j = \sum_{p,q=1}^r (\bar{x}_p^j - \bar{x}_q^j)^2$, $b^j = \sum_{p=1}^r \frac{1}{m_p} \sum_{i=1}^{m_p} (\bar{x}_{pi}^j - \bar{x}_{qi}^j)^2$, $j = \overline{1, N}$.

Тогда функционал (4) сводится к виду (5). Критерий этого вида обозначим через $I(\lambda)$

$$I(\lambda) = \frac{(a, \lambda)}{(b, \lambda)}, \quad (5)$$

где $(*, *)$ - скалярное произведение векторов.

Коэффициенты a^j, b^j не зависят от λ и вычисляются заранее. Для расчета функционала $I(\lambda)$ при каждом λ требуется порядка N операций.

Определение 3. Если для $\forall \alpha \in R, (\alpha \neq 0)$ выполняется равенство $f(\alpha x_1, \alpha x_2, \dots, \alpha x_k) = \alpha^k f(x_1, x_2, \dots, x_k)$, то f называется однородной функцией k -порядка.

Учитывая данное определение, в параграфе 2.1 приводятся критерии формирования информативного пространства признаков, представленные однородными функционалами 0-порядка обобщенного вида

$$I(\lambda) = \frac{(a, \lambda)^r}{\prod_{j=1}^r (b^{(j)}, \lambda)}; \quad (6)$$

$$I(\lambda) = \prod_{j=1}^r \frac{(a^{(j)}, \lambda)}{(b^{(j)}, \lambda)}. \quad (7)$$

Критерии, представленные функционалами (6) и (7), целесообразно использовать, когда количество заданных классов больше или равно двум. Кроме таких критериев существуют критерии, определяемые однородными функционалами k -порядка, а также неоднородными функционалами. Так как эти функционалы в проблематике распознавания образов являются малоисследованными, то в настоящее время отсутствуют методы и алгоритмы выбора информативных признаков, основанные на использовании критериев, задаваемых подобными функционалами.

Пользуясь приведенными выше обозначениями и определениями, рассмотрим критерий, заданный функционалом

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

где c – некоторый вектор внешних факторов. Например, каждая компонента данного вектора может представлять средства, требуемые для измерения соответствующего признака.

Функционал (8) является однородным “-1”-порядка. При выборе информативных признаков решается задача максимизации функционала (8).

В случае, если количество классов равно трем ($r=3$), то можно использовать критерий, заданный функционалом

$$I(\lambda) = \frac{(a, \lambda)}{(b, \lambda)(c, \lambda)(d, \lambda)}. \quad (9)$$

В общем виде однородный функционал k -порядка можно представить одним из следующих выражений:

$$I(\lambda) = \frac{\prod_{j=1}^r (a^{(j)}, \lambda)}{(b, \lambda)}, \quad (10)$$

$$I(\lambda) = \frac{(a, \lambda)}{\prod_{i=1}^r (b^{(i)}, \lambda)}. \quad (11)$$

Функционал (10) является положительным однородным функционалом k -порядка, а (11) – отрицательным. Критерий, заданный однородным функционалом k -порядка, можно представить как

$$I(\lambda) = \frac{\prod_{j=1}^t (a^{(j)}, \lambda)}{\prod_{i=1}^h (b^{(i)}, \lambda)}, \quad (t, h \in N). \quad (12)$$

Необходимо отметить, что критерии, заданные однородными функционалами k -порядка, считаются более эффективными, чем критерии заданные однородными функционалами нулевого порядка, так как они учитывают закономерности расположения объектов внутри каждого класса и не изменяют расстояние между классами.

При решении практических задач кроме однородных критериев также используются критерии, заданные неоднородными функционалами, например,

$$I(\lambda) = \frac{(a, \lambda)}{(b, \lambda)} + \frac{(a^*, \lambda)}{(b^*, \lambda)(c^*, \lambda)}. \quad (13)$$

Нетрудно увидеть, что функционал (13) представляет собой сумму однородных функционалов нулевого и k -порядка.

В общем виде неоднородный критерий фишеровского типа формируется комбинацией критериев, заданных функционалами (6)-(8), (10)-(12).

В параграфе 2.2 диссертации приводятся некоторые особенности критерия информативности, представленного функционалом (5).

Пусть Λ^ℓ -множество всех ℓ -информативных векторов. Из определения ℓ -информативного вектора следует, что мощность множества Λ^ℓ равна C_N^ℓ .

Введем функцию

$$F(\ell) = \max_{\lambda \in \Lambda^\ell} I(\lambda). \quad (14)$$

При заданном ℓ требуется найти такой ℓ -информативный вектор $\lambda_\ell \in \Lambda^\ell$, при котором имеет место равенство $I(\lambda_\ell) = \max_{\lambda \in \Lambda^\ell} I(\lambda)$.

Решение этой задачи осуществляется с помощью лемм 1-6, приведённых в параграфе 2.2, а также следующих теорем, приведенных в том же параграфе.

Пусть даны действительные числа a, b и $c \geq 0, d > 0$ ($a+c \geq 0, b+d > 0$).

Тогда имеет место одна из следующих лемм:

Лемма 1. Если $\begin{cases} a > 0 \\ b > 0 \end{cases}$ и $\frac{c}{d} > \frac{a}{b}$, то имеет место соотношение $\frac{a}{b} < \frac{a+c}{b+d} < \frac{c}{d}$.

Лемма 2. Если $\begin{cases} a > 0 \\ b > 0 \end{cases}$ и $\frac{c}{d} < \frac{a}{b}$, то имеет место соотношение $\frac{a}{b} > \frac{a+c}{b+d} > \frac{c}{d}$.

Лемма 3. Если $\begin{cases} a < 0 \\ b < 0 \end{cases}$ и $\frac{c}{d} < \frac{a}{b}$, то имеет место соотношение $\frac{a}{b} > \frac{a+c}{b+d} < \frac{c}{d}$.

Лемма 4. Если $\begin{cases} a < 0 \\ b < 0 \end{cases}$ и $\frac{c}{d} > \frac{a}{b}$, то имеет место соотношение $\frac{a}{b} < \frac{a+c}{b+d} > \frac{c}{d}$.

Лемма 5. Если $\begin{cases} a \geq 0 \\ b \leq 0 \end{cases}$, то имеет место соотношение $\frac{a+c}{b+d} \geq \frac{c}{d}$.

Лемма 6. Если $\begin{cases} a \leq 0 \\ b \geq 0 \end{cases}$, то имеет место соотношение $\frac{a+c}{b+d} \leq \frac{c}{d}$.

Теорема 1. Для $\forall l (l > 0)$ $\max\{I(\lambda) : \lambda \in \Lambda^\ell\} = \frac{a^1 + \sum_{k=1}^{\ell-1} a^{j_k}}{b_1 + \sum_{k=1}^{\ell-1} b^{j_k}}$.

Для функции (14) справедлива следующая теорема.

Теорема 2. Функция $F(\ell)$ не возрастает при увеличении аргумента ℓ .

Исследование свойств критерия информативности Фишера (5) показало, что полученные результаты можно использовать в качестве основы для построения метода выбора информативных наборов признаков.

В третьей главе диссертации “Методы формирования информативного описания объектов с использованием эвристических критериев информативности признаков” предложены методы и соответствующие им алгоритмы определения информативных наборов признаков в задачах распознавания образов.

В параграфе 3.1 предлагаются регулярный метод, рекуррентный метод, метод частичного перебора, метод упорядочения и метод "Дельта" для определения информативных наборов признаков с использованием эвристического критерия, заданного функционалом (5). При этом решаемая оптимизационная задача формулируется как

$$\begin{cases} I(\lambda) \rightarrow \max; \\ \lambda \in \Lambda^\ell, \end{cases} \quad (15)$$

где Λ^ℓ - множество всех ℓ - информативных признаков.

Для решения этой задачи с использованием регулярного метода введем вектор-функцию

$$\varphi(\lambda) = a(b, \lambda) - b(a, \lambda), \quad (16)$$

которая указывает направление наискорейшего роста функционала $I(\lambda)$ в точке λ . Этот метод основывается на следующих теоремах и свойстве.

Пусть $\lambda \in \Lambda^\ell$ выбранный вектор. При $\forall \mu \in \Lambda^\ell$ справедлива следующая теорема.

Теорема 3. Для того, чтобы выполнялось $I(\lambda) > I(\mu)$, необходимо и достаточно отсутствие вектора μ , удовлетворяющего условию $(\varphi(\lambda), \mu) > 0$.

Свойство 1. Для произвольного $\lambda (\lambda \in \Lambda^\ell)$ верно $(\varphi(\lambda), \mu(\lambda)) \geq 0$.

Из теоремы 3 и свойства 1 вытекает $I(\lambda) \leq I(\mu(\lambda))$.

Теорема 4. Если $I(\lambda) = I(\mu(\lambda))$, то $I(\lambda) = \max \{I(\eta) | \eta \in \Lambda^\ell\}$.

Данная теорема гарантирует оптимальность полученного решения, т.е. значения функционала $I(\lambda)$ при найденном решении λ достигают своего максимума на множестве Λ^ℓ .

На теоремах 3 и 4 основан регулярный метод, реализуемый в виде итерационной процедуры максимизации функционала (5). При этом на первом шаге выбирается произвольный ℓ - информативный вектор λ , например, $\lambda = (\overbrace{1, 1, \dots, 1}^{\ell}, 0, 0, \dots, 0)$. Далее на каждой итерации новый вектор λ определяется из предыдущего с помощью оператора следования $\mu(\lambda)$, т.е. $\lambda = \mu(\lambda)$. Итерационный процесс продолжается до тех пор, пока функционал $I(\lambda)$ растет. В случае, если рост прекращается, т.е. $I(\lambda) = I(\mu(\lambda))$, то λ - оптимальное решение. Обычно это решение достигается, как показали эксперименты, на 3-4 шаге.

Цель метода частичного перебора, предложенный в диссертации, заключается в нахождении ℓ - информативного вектора λ , обеспечивающего решение задачи (15). При этом количество вычислений функционала $I(\lambda)$ для заданного $\ell \geq 2$ равно

$$\frac{\prod_{i=1}^{\ell-1} (N - \ell + i)}{\ell!}.$$

Суть предложенного в диссертации метода упорядочения для выбора информативного набора признаков на основе решения задачи (15) состоит в следующем.

Пусть, отношения соответствующих компонентов векторов a и b упорядочены следующим образом:

$$\frac{a_1}{b_1} \geq \frac{a_2}{b_2} \geq \dots \geq \frac{a_N}{b_N}. \quad (17)$$

Тогда вектор $\lambda = (1, 1, \dots, 1, \underbrace{0, 0, \dots, 0}_{N-l})$ является оптимальным решением задачи (15) по методу упорядочения. Однако, этот метод не всегда обеспечивает оптимальное решение этой задачи по критерию (5). Например, легко можно показать, что для $a = (5, 10, 10, 1)$, $b = (1, 50, 50, 19)$ и $N = 4$, $\ell = 2$ оптимальным решением является вектор $\lambda = (1, 0, 0, 1)$, а не вектор $\lambda = (1, 1, 0, 0)$, полученный с использованием метода упорядочения.

Ниже приведены условия, при выполнении которых метод упорядочения обеспечивает оптимальное решение задачи (15).

Введем следующие обозначения:

$$A = \sum_{i=1}^l a_i, \quad B = \sum_{i=1}^l b_i, \quad \begin{cases} \Delta a_{ij} = a_j - a_i \\ \Delta b_{ij} = b_j - b_i, i = \overline{1, l}, j = \overline{l+1, N} \end{cases}, \quad \lambda^0 = \left(1, 1, \dots, 1, \underbrace{0, 0, \dots, 0}_{N-l} \right).$$

Если в леммах, приведенных во главе 2, принято

$$a = \Delta a_{ij}, \quad b = \Delta b_{ij}, \quad c = A, \quad d = B,$$

то для $\forall i, j (i = \overline{1, \ell}, j = \overline{\ell+1, N})$, с учетом $\begin{cases} A + \Delta a_{ij} \geq 0, \\ B + \Delta b_{ij} > 0 \end{cases}$, будет иметь место одна из лемм 1-6.

Теорема 5. Для того, чтобы вектор $\lambda^0 = \left(1, 1, \dots, 1, \underbrace{0, 0, \dots, 0}_{N-l} \right)$, выбранный с помощью упорядоченной последовательности (17), был оптимальным решением задачи (15), необходимо и достаточно отсутствие соотношений $a = \Delta a_{ij}$, $b = \Delta b_{ij}$ в условиях лемм 2 и 4.

Метод «Дельта», разработанный на основе метода упорядочения, реализуется следующим образом.

Допустим, $\lambda = \left(\underbrace{1, 1, 1, \dots, 1}_{\ell}, \underbrace{0, 0, 0, \dots, 0}_{N-\ell} \right)$ является вектором, полученным по методу упорядочения. Предположим, что этот вектор не является оптимальным решением задачи (15). Тогда для нахождения такого решения необходимо осуществить такие замены на основе лемм 2 и 4, пока не будут исчерпаны все Δa_{ij} и Δb_{ij} , удовлетворяющие условиям этих лемм. При этом, в соответствии с теоремой 5, найденное решение будет оптимальным.

Во многих случаях предварительно выбранный вектор λ может обеспечить оптимальное решение задачи (15), поэтому следующая теорема позволяет определить, при каких условиях это может произойти.

Пусть выбран $\forall \lambda \in \Lambda'$.

Теорема 6. Для того, чтобы выбранный вектор λ обеспечил оптимальное решение задачи (15), необходимо и достаточно отсутствия $a = \Delta a_{ij}$ и $b = \Delta b_{ij}$ ($i = \overline{1, l}, j = \overline{l+1, N}$), удовлетворяющих условиям лемм 2, 4 и 5.

В методе «Дельта» значения функционала и компоненты вектора λ определяются как и в методе упорядочения.

В параграфе 3.2 диссертации рассмотрено решение задачи (15) на основе критерия, задаваемого в виде функционала (6).

Предварительно вектор λ выберем в виде $\lambda^0 = (\underbrace{1, 1, \dots, 1}_{\ell}, \underbrace{0, 0, \dots, 0}_{N-\ell})$ и введем обозначения

$$A = \sum_{i=1}^l a_i, \quad B_j = \sum_{i=1}^l b_i^j, \quad j = \overline{1, r},$$

а также введем вспомогательный вектор C , компоненты которого вычисляются следующим образом: $c_i = \frac{a_i}{A} - \sum_{j=1}^r \frac{b_i^{(j)}}{B_j}$.

Допустим, дан $\forall \mu \in \Lambda^\ell$ и в соответствии с λ^0 сформирован вектор C .

Теорема 7. Для того, чтобы выбранный вектор λ обеспечил оптимальное решение задачи (15), необходимо и достаточно отсутствие вектора μ , удовлетворяющего условию $(C, \mu) > 0$.

На этой теореме основан метод, названный в диссертации как «Градиент-1». Вектор μ в этом методе определяется также как и в регулярном методе.

Пусть дан критерий (8), который является однородным критерием «-1» порядка. Рассмотрим задачу (15).

Введем следующие обозначения:

$$A = \sum_{i=1}^l a_i, \quad B = \sum_{i=1}^l b_i, \quad C = \sum_{i=1}^l c_i, \quad \begin{cases} \Delta a_{ij} = a_j - a_i, \\ \Delta b_{ij} = b_j - b_i, i = \overline{1, l}, j = \overline{l+1, N}. \end{cases}$$

Пусть даны действительные числа a, b, c , а также $d, e, f > 0$, удовлетворяющие условиям $b+e > 0, c+f > 0$.

Лемма 7. Если $\frac{a}{d} \geq \frac{b}{e} + \frac{c}{f} + \frac{bc}{ef}$, тогда имеет место $\frac{d}{ef} \leq \frac{d+a}{(b+e)(c+f)}$.

Лемма 8. Если $\frac{a}{d} < \frac{b}{e} + \frac{c}{f} + \frac{bc}{ef}$, тогда имеет место $\frac{d}{ef} > \frac{d+a}{(b+e)(c+f)}$.

Пусть выбран $\forall \lambda \in \Lambda^\ell$.

Теорема 8. Для того, чтобы выбранный вектор λ обеспечил оптимальное решение задачи (15), необходимо и достаточно отсутствие $a = \Delta a_{ij}, b = \Delta b_{ij}$ и $c = \Delta c_{ij}$ ($i = \overline{1, l}, j = \overline{l+1, N}$), удовлетворяющих условиям леммы 7.

В данном методе значения функционала и компоненты вектора λ определяются как в методе упорядочения.

Кроме того, в параграфе 3.2 диссертации рассмотрено решение задачи (15) для критерия, заданного в виде однородного функционала с положительной степенью (10).

В целях упрощения вычислений введем обозначения

$$A^{(j)} = (a^{(j)}, \lambda), B = (b, \lambda), j = \overline{1, r}, A_1^{(j)} = (a^{(j)}, \mu), B_1 = (b, \mu),$$

а также вектор $C = (c_1, c_2, \dots, c_N)$, компоненты которого определяются

$$\text{следующим образом: } c_i = \frac{\sum_{t=1}^r a_i^{(t)}}{w^{r-1}} - I_\lambda b_i, i = \overline{1, N}.$$

Пусть заданы ℓ -информативные векторы λ и μ , а также $b_i > 0, i = \overline{1, N}$. Тогда для $I(\lambda)$, $I(\mu)$ и сформированного вектора C имеет место следующая теорема.

Теорема 9. Для того, чтобы имело место неравенство $I(\lambda) < I(\mu)$, достаточно, чтобы $(C, \mu) > 0$.

На этой теореме основан метод, названный в диссертации как «Градиент-2». Вектор μ в этом методе определяется также как и в регулярном методе. При этом данный метод можно использовать для выбора информативных наборов признаков на основе однородных критериев отрицательного $1-r$ порядка.

Функционал (11) выражает общий вид однородного критерия фишерского типа с отрицательной степенью k ($k \leq 0$).

Для решения задачи (15) введем градиентный вектор $C = (c_1, c_2, \dots, c_N)$. Компоненты этого вектора вычисляются с помощью следующей формулы

$$c_i = \frac{a_i}{w^{r-1}} - I_\lambda \sum_{j=1}^r b_i^{(j)}, i = \overline{1, N}. \text{ Для вектора } C = (c_1, c_2, \dots, c_N) \text{ справедлива теорема 9.}$$

Метод «Градиент-2» можно использовать также для критерия, представленного в виде комбинации из критериев фишерского типа. При этом направляющий вектор формируется на основе комбинации направляющих векторов, соответствующих исходным критериям. Для комбинированного направляющего вектора также справедлива теорема 9.

В четвертой главе диссертации «**Методы определения неинформативных наборов признаков**» изложены методы определения неинформативных и малоинформативных наборов признаков на основе рассмотренных в диссертации эвристических критериев информативности.

Множество $R^N \setminus \Lambda^\ell$ называется неинформативным и обозначается как \aleph , т.е. $\aleph = R^N \setminus \Lambda^\ell$. Здесь исходное пространство признаков выражается в виде $\{\Lambda^\ell, \aleph\}$, т.е. $R^N = \Lambda^\ell \cup \aleph$.

В настоящее время большая часть существующих методов направлена на определение информативных признаков. Это так называемые «прямые» методы. Однако существуют также методы определения информативных признаков путем выявления и последующего исключения неинформативных признаков.

Пусть Λ^ℓ -множество информативных признаков, полученных при выборе информативных признаков, а \aleph -множество неинформативных признаков, полученных при выборе неинформативных признаков. Тогда:

1. Если $\Lambda^\ell \cap \aleph = \emptyset$, тогда $\Lambda^\ell \cup \aleph = R^N$, $\dim(\Lambda^\ell) + \dim(\aleph) = N$.
2. Если $K = \Lambda^\ell \cap \aleph \neq \emptyset$, тогда $\Lambda^\ell \cup \aleph \subset R^N$, $\Lambda^\ell \setminus K$ называется множеством информативных признаков, $\aleph \setminus K$ - множеством неинформативных признаков,

а K - множеством малоинформативных признаков. где $\dim(\Lambda^\ell) + \dim(\aleph) > N$, $\dim(\Lambda^\ell \setminus K) + \dim(\aleph \setminus K) + \dim(K) = N$ и исходное пространство признаков выражается как $\{\Lambda^\ell, K, \aleph\}$, т.е. $R^N = \Lambda^\ell \cup K \cup \aleph$.

Предлагаемые в диссертации методы выбора неинформативных наборов признаков являются модификациями методов выбора информативных наборов признаков, приведенных в параграфе 3.1.

Модифицированный регулярный метод выбора неинформативных наборов признаков основан на использовании критерия информативности, заданного функционалом (5). При этом решаемая задача формулируется как

$$\begin{cases} I(\eta) \rightarrow \min; \\ \eta \in \aleph^{\wp}, \end{cases} \quad (18)$$

где \aleph^{\wp} - множество всех \wp -информативных признаков.

Для решения задачи (18) введем вектор-функцию $\varphi(\lambda) = a(b, \lambda) - b(a, \lambda)$, которая указывает направление наискорейшего уменьшения функционала $I(\eta)$ в точке η .

Теорема 10. Для того, чтобы выполнялось неравенство $I(\eta) < I(\mu)$, необходимо и достаточно отсутствие \wp -неинформативного вектора μ , удовлетворяющего условию $(\varphi(\eta), \mu) < 0$.

Введем оператор (следования): $\mu: \aleph^{\wp} \rightarrow \aleph^{\wp}$. Здесь $(\varphi(\eta), \mu(\eta)) = \min_{\vartheta \in \aleph^{\wp}} (\varphi(\eta), \vartheta)$.

Оператор μ имеет очевидное конструктивное представление, если компоненты вектора $\varphi(\eta)$ упорядочены в порядке убывания, т.е. $\varphi^{j_1}(\eta) \leq \varphi^{j_2}(\eta) \leq \dots \leq \varphi^{j_N}(\eta)$, тогда компоненты векторы $\mu(\eta)$ будут определены как $\mu^{j_1}(\eta) = 1, \mu^{j_2}(\eta) = 1, \dots, \mu^{j_{\ell}}(\eta) = 1, \mu^{j_{\ell+1}}(\eta) = 0, \mu^{j_{\ell+2}}(\eta) = 0, \dots, \mu^{j_N}(\eta) = 0$.

Иначе говоря, компоненты вектора $\varphi(\eta)$, соответствующие первым η -максимальным компонентам вектора $\mu(\eta)$, равны единице, а остальные - нулю.

Очевидно, что вектор $\mu(\eta)$ является \wp -неинформативным вектором, для которого имеет место следующее условие.

$$(\varphi(\eta), \mu(\lambda)) = \min \{(\varphi(\eta), \vartheta) \mid \vartheta \in \aleph^{\wp}\}. \quad (19)$$

Свойство 2. Для произвольного $\eta (\eta \in \aleph^{\wp})$ верно $(\varphi(\eta), \mu(\eta)) \leq 0$.

Теорема 11. Если $I(\eta) = I(\mu(\eta))$, то $I(\eta) = \min \{I(\vartheta) \mid \vartheta \in \aleph^{\wp}\}$.

Эта теорема гарантирует оптимальность полученного решения, т.е. функционал $I(\eta)$ при найденном решении η достигает своего минимума на множестве \aleph^{\wp} .

Метод упорядочения, используемый для определения информативных признаков на основе критерия (5), можно модифицировать для поиска неинформативных признаков на основе решения следующей оптимизационной задачи.

Пусть отношения соответствующих компонентов векторов a и b упорядочены следующим образом:

$$\frac{a_1}{b_1} \leq \frac{a_2}{b_2} \leq \dots \leq \frac{a_N}{b_N}. \quad (20)$$

Тогда вектор $\eta = (\underbrace{1,1,1,\dots,1}_\varphi, \underbrace{0,0,\dots,0}_{N-\varphi})$ является оптимальным решением задачи (18) по методу упорядочения. Однако этот метод не всегда обеспечивает оптимальное решение по отношению критерия Фишера.

Введем следующие обозначения:

$$A = \sum_{i=1}^l a_i, \quad B = \sum_{i=1}^l b_i, \quad \begin{cases} \Delta a_{ij} = a_j - a_i \\ \Delta b_{ij} = b_j - b_i, i = \overline{1, l}, j = \overline{l+1, N} \end{cases}, \quad \eta^0 = \left(\underbrace{1,1,\dots,1}_\varphi, \underbrace{0,0,\dots,0}_{N-\varphi} \right).$$

Если в леммах 1-6, приведенных выше, принято $a = \Delta a_{ij}, b = \Delta b_{ij}, c = A, d = B$, то для $\forall i, j (i = \overline{1, l}, j = \overline{\varphi+1, N})$, с учетом

$$\begin{cases} A + \Delta a_{ij} \geq 0; \\ B + \Delta b_{ij} > 0, \end{cases}$$

будет иметь место одна из этих лемм.

Теорема 12. Для того, чтобы вектор $\eta^0 = (\underbrace{1,1,\dots,1}_\varphi, \underbrace{0,0,\dots,0}_{N-\varphi})$, выбранный с помощью упорядоченной последовательности (20), был оптимальным решением задачи (18), необходимо и достаточно отсутствие соотношений $a = \Delta a_{ij}, b = \Delta b_{ij}$ в условиях лемм 1 и 3.

Модифицированный метод «Дельта», ориентированный на выбор неинформативных наборов признаков с использованием критерия (5), основан на следующей теореме.

Пусть выбран $\forall \eta \in \mathcal{N}^\varphi$.

Теорема 13. Для того, чтобы выбранный вектор η обеспечил оптимальное решение задачи (18), необходимо и достаточно отсутствие $a = \Delta a_{ij}$ и $b = \Delta b_{ij} (i = \overline{1, l}, j = \overline{l+1, N})$, удовлетворяющих условиям лемм 1, 3 и 6.

Если вектор η не является оптимальным решением задачи (18), то осуществляются замены на основе лемм 1, 3 и 6. Процесс замены продолжается до тех пор, пока не будут исчерпаны все Δa_{ij} и Δb_{ij} , удовлетворяющие условиям лемм 1, 3 и 6. При этом, в соответствии с теоремой 13, найденное решение будет оптимальным.

В параграфе 4.2 приведены методы выбора неинформативного набора признаков по критерию, заданному в виде обобщенного однородного функционала нулевого порядка (6).

Вектор η предварительно выберем в виде $\eta^0 = (\underbrace{1,1,\dots,1}_\varphi, \underbrace{0,0,\dots,0}_{N-\varphi})$ и введем

следующие обозначения: $A = \sum_{i=1}^l a_i, B_j = \sum_{i=1}^l b_i^j, j = \overline{1, r}$.

Для решения задачи (18) относительно функционала (6) введем вспомогательный вектор C , компоненты которого вычисляются по формуле

$$c_i = \frac{a_i}{A} - \sum_{j=1}^r \frac{b_i^{(j)}}{B_j}.$$

Допустим, дан $\forall \mu \in \mathbb{N}^{\rho}$ и в соответствии с η^0 сформирован вектор C .

Теорема 14. Для того чтобы выбранный вектор η обеспечил оптимальное решение задачи (18), необходимо и достаточно отсутствие вектора μ , удовлетворяющего условию $(C, \mu) < 0$.

Метод «Градиент-1», ориентированный на выбор неинформативных наборов признаков, основан на теореме 14 и сводится к реализации процесса формирования вектора μ , который длится до выполнения условия $I(\eta) = I(\mu)$.

Рассмотрим оптимизационную задачу (18), решаемую относительно функционала (8).

Введем следующие обозначения:

$$A = \sum_{i=1}^l a_i, \quad B = \sum_{i=1}^l b_i, \quad C = \sum_{i=1}^l c_i, \quad \begin{cases} \Delta a_{ij} = a_j - a_i, \\ \Delta b_{ij} = b_j - b_i, \quad i = \overline{1, l}, \quad j = \overline{l+1, N}. \end{cases}$$

Пусть выбран $\forall \eta \in \mathbb{N}^{\rho}$.

Теорема 15. Для того, чтобы выбранный вектор η обеспечил оптимальное решение задачи (18), необходимо и достаточно отсутствие $a = \Delta a_{ij}, b = \Delta b_{ij}$ и $c = \Delta c_{ij}$ ($i = \overline{1, l}, j = \overline{l+1, N}$), удовлетворяющих условиям леммы 8.

Модифицированный метод «Дельта», используемый для выбора неинформативных признаков по критерию (8), основан на теореме 15 и реализуется аналогично методу упорядочения.

Кроме того, в параграфе 4.2 рассмотрено применение метода «Градиент-2» для решения задачи (18) относительно функционала (8).

Введем вспомогательный вектор $D(d_1, d_2, \dots, d_n)$, координаты которого определяются следующим образом: $d_i = \frac{a_i}{A} - \frac{b_i}{B} - \frac{c_i}{C}$. Здесь

$$A = (a, \eta), \quad B = (b, \eta), \quad D = (d, \eta), \quad A_1 = (a, \mu), \quad B_1 = (b, \mu), \quad D_1 = (d, \mu).$$

Тогда соответствующие значения функционала (8) равны $I(\eta) = \frac{A}{BD}$ и $I(\mu) = \frac{A_1}{B_1 D_1}$.

Для выбранного вектора λ и сформированного на его основе вектора C справедлива следующая теорема.

Теорема 16. Для того, чтобы выполнялось неравенство $I(\eta) < I(\mu)$, достаточно, чтобы $(C, \mu) > 0$.

Теорема 17. Если $I(\eta) = I(\mu)$, то $I(\eta) = \min \{I(\mathcal{G}) \mid \mathcal{G} \in \mathbb{N}^{\rho}\}$.

Последняя теорема обеспечивает оптимальность решения, выбранного из области \mathbb{N}^{ρ} .

С помощью теорем 16 и 17 обосновывается итерационный метод, оптимизирующий функционал (11). На первом шаге этого метода произвольно выбирается ρ -неинформативный вектор η . Для удобства вектор η можно выбрать в виде $\eta^0 = (\overbrace{1, 1, \dots, 1}^{\rho}, 0, 0, \dots, 0)$. В последующем, на каждом шаге новый вектор η формируется на основе построения предыдущего

вектора μ . Итерационный процесс продолжается до тех пор, пока не будет выполнено $(C, \eta) > 0$.

Однородные или неоднородные смешанные критерии k -порядка представляются в виде комбинации критериев нулевого и k -порядка и в этом случае метод формирования пространства неинформативных признаков осуществляется аналогично методу выбора неинформативных признаков “Градиент-2”, где направляющий вектор формируется на основе комбинации направляющих векторов, соответствующих критериям нулевого и k -порядка.

Пятая глава диссертации **“Экспериментальные исследования и практическое применение”** отражает методику формирования признакового пространства, а также обучающей и контрольной выборок, используемых для проведения экспериментальных исследований предложенных в диссертации методов и алгоритмов выбора информативных признаков. Кроме того, в данной главе приводятся описание и результаты практического применения программного обеспечения “MDIB”.

В параграфе 5.1 сформирована исходная база данных, включающая обучающие и контрольные выборки, которые отвечают следующим требованиям:

- обучающие и контрольные выборки создаются на основе одинаковых законов и правил;
- количество классов объектов обучающей и контрольной выборок должно быть не менее трех;
- заранее известны ℓ информативные признаки ($1 \leq \ell \leq N$).

Параграф 5.2 отражает результаты экспериментальных исследований, которые проводились для трех классов объектов K_1, K_2 и K_3 , заданных в R^N признаковом пространстве. Размерность N выбрана кратной некоторому заданному небольшому числу ℓ , т.е. $N = \ell * h$ ($h \in N$). Набор признаков $x = (x_1, x_2, \dots, x_N)$ скомпонован по группам $(x_1, \dots, x_\ell, x_{\ell+1}, \dots, x_{2\ell}, x_{2\ell+1}, \dots, x_{(h-1)\ell+1}, \dots, x_{h\ell})$, каждая из которых содержит ℓ признаков. Классы выбраны так, что с точки зрения их разделимости группы признаков взаимно независимы, а внутри одной группы признаки сильно зависимы, т.е. удаление любого признака из своей группы влечет за собой резкое ухудшение информативности этой группы в целом.

Эксперименты были связаны с решением оптимизационной задачи для функционала (8) с использованием методов “Дельта-2” и полного перебора. Результаты определения информативных наборов признаков, полученные с помощью метода “Дельта-2”, полностью совпали с результатами, полученными на основе метода полного перебора. При этом число итераций при полном переборе составило C_N^ℓ , а для метода “Дельта-2” максимальное число итераций было равно ℓ .

Параграф 5.3 диссертационной работы посвящен описанию программного обеспечения “MDIB”, созданного на основе существующих и предложенных в диссертации алгоритмов предварительной обработки данных.

Параграф 5.4 отражает результаты практического применения программного обеспечения “MDIB” для решения задачи обеспечения безопасной работы комплекса КНК-270 на основе предварительной обработки данных, получаемых от датчиков, установленных на этом комплексе.

В рабочем режиме комплекс КНК-270, установленный в Навоинском горно-металлургическом комбинате, может находиться в одном из трех состояний: опасное, промежуточное, безопасное. Сигналы, характеризующие опасное состояние комплекса, представляли собой первый класс объектов K_1 , сигналы, характеризующие промежуточное состояние - второй класс объектов K_2 , сигналы, характеризующие безопасное состояние - третий класс объектов K_3 .

Число признаков, характеризующих каждый объект, равно 9.

Каждый класс содержит одинаковое число объектов, равное 5056. Таким образом, каждый объект можно представить в виде вектора $x_{ij} = (x_{ij}^1, x_{ij}^2, \dots, x_{ij}^9)$, где x_{ij}^k - k -ый признак j -го объекта i -го класса, где $k = \overline{1,9}$; $j = \overline{1,5056}$; $i = \overline{1,3}$.

Математическая постановка задачи определения основных показателей, характеризующих состояния комплекса КНК-270 заключалась в следующем:

$$\begin{cases} I(\lambda) = \frac{(a, \lambda)}{(b, \lambda)(c, \lambda)(d, \lambda)} \rightarrow opt, \\ \lambda \in \Lambda^\ell. \end{cases}$$

После определения информативных наборов признаков (для $\ell = \overline{1,9}$) была решена задача распознавания с использованием метода “ k -ближайших соседей”, что позволило оценить степень «полезности» каждого из этих наборов признаков с точки зрения их влияния на качество распознавания объектов контрольной выборки.

В результате решения этой задачи был определен наиболее информативный набор признаков x_2, x_5, x_8 , где x_2, x_5 и x_8 представляют значения сигналов, поступающих от вертикальных компонент трехкомпонентных датчиков $D_i, i = \overline{1,3}$.

На основе этих результатов специалистами предметной области были разработаны рекомендации по предотвращению аварийных ситуаций в процессе эксплуатации комплекса КНК-270.

ЗАКЛЮЧЕНИЕ

Результаты проведенного диссертационного исследования на тему «Методы и алгоритмы предварительной обработки данных» сводятся к следующим основным выводам:

1. Формализована проблема снижения размерности в задачах анализа данных. Это служит формированию задачи выбора информативных признаков. Разработаны эвристические критерии информативности и неинформативности признаков. Эти критерии дают возможность выявить основные характеризующие показатели объекта.

2. Разработаны методы и алгоритмы выбора наборов информативных и неинформативных признаков на основе единого критерия информативности. Настоящие методы и алгоритмы служат повышению достоверности информативных признаков.

3. Задача снижения размерности исходного пространства признаков приведена к оптимизационной задаче, которая заключается в нахождении в N -мерной исходной системе признаков, такой l -мерной ($l < N$) подсистемы признаков, при которой обеспечивается экстремальное значения заданной меры информативности признаков. Решение настоящей оптимизационной задачи служит информативному описанию объектов при наименьших количествах признаков.

4. На основе анализа свойств эвристических критериев информативности Фишера и Горелика выявлено принципиальное отличие этих двух критериев. При этом функционал Горелика не производит суммирование частных критериев, а выбирает лишь один, который является максимальным или минимальным. И это служит повышению достоверности разделимости классов.

5. Выявлены особенности критериев информативности Фишера, которые в дальнейшем использованы в качестве конструктивной основы для разработки методов выбора информативных наборов признаков. В частности, аналитически определены верхняя и нижняя границы критерия Фишера для заданной l -информативной подсистемы признаков. Показано, что признак, определяющий максимальное значение отношения соответствующих компонент вектора, характеризующего межклассовые расстояния, и вектора, характеризующего внутриклассовой разброс объектов, всегда присутствует в l -информативной подсистеме признаков, обеспечивающей максимальное значение критерия информативности. Это свойство критерия Фишера обуславливает возможность перехода от N -мерного признакового пространства к $N-l$ -мерному пространству.

6. Для критерия информативности Фишера предложены методы выбора информативных признаков, названные в работе регулярным, рекуррентным и методом частичного перебора. Показано, что для решения оптимизационной задачи выбора информативных признаков регулярный метод использует вектор-функцию, которая указывает направление наискорейшего роста функционала $I(\lambda)$. Для случая выполнения условия

упорядочения разработан рекуррентный метод, который позволяет, в отличие от регулярного метода, за фиксированное количество $((2N-l+1)/2)$ шагов получить оптимальное решение. Это даёт возможность решения оптимизационной задачи методом частичного перебора.

7. Показано, что метод упорядочения не всегда обеспечивает оптимальное решение оптимизационной задачи на основе информативных критериев и определены оптимальные условия. Эти условия дают возможность оптимальное решение из субоптимальных решений.

8. Определены необходимые и достаточные оптимальные условия выбранного решения и на основе этих решений разработан ускоренный метод Дельта. Метод Дельта служит определению оптимальности решений, полученных при формировании пространства информативных признаков на основе простых критериев фишеровского типа, также при получении оптимального решения из субоптимального решения, полученного с помощью других методов.

9. Разработаны методы и алгоритмы определения информативных и неинформативных наборов признаков с использованием обобщенных критериев информативности нулевого порядка. Результаты, полученные на основе этих методов и алгоритмов, служат информативному описанию объектов.

10. Разработаны методы и алгоритмы определения информативных и неинформативных наборов признаков с использованием обобщенных критериев информативности k -го порядка. Настоящие методы и алгоритмы, учитывая закономерности расположения объектов между каждым классом, дают возможность информативного описания объектов.

11. Разработаны методы и алгоритмы определения информативных и неинформативных наборов признаков с использованием критериев информативности смешанного типа. Разработанные методы и алгоритмы дают возможность информативному описанию объектов при различных размерностях признакового пространства.

12. Предложены методы и алгоритмы определения неинформативных наборов признаков, реализация которых связана с минимизацией заданного эвристического критерия информативности. Принципы построения этих методов по сути близки к принципам, на которых базируются методы определения информативных наборов признаков. Отличие между ними обусловлено лишь тем, что вместо задачи максимизации критерия информативности (при определении информативных наборов признаков) решается задача минимизации (при определении неинформативных наборов признаков).

**AD HOC SCIENTIFIC COUNCIL AT THE SCIENTIFIC COUNCIL
AWARDING SCIENTIFIC DEGREES DSc.27.06.2017.T.07.01 AT
TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES**

**SCIENTIFIC AND INNOVATION CENTER OF INFORMATION AND
COMMUNICATION TECHNOLOGIES AT THE TASHKENT
UNIVERSITY OF INFORMATION TECHNOLOGIES
NAMED AFTER MUHAMMAD AL-KHWARIZMI**

MAMATOV NARZILLO SOLIDJONOVICH

**METHODS AND ALGORITHMS OF PRELIMINARY DATA
PROCESSING**

05.01.03 – Theoretical basis of computer science

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

Tashkent-2018

The theme of doctoral dissertation of technical sciences (DSc) was registered with the number of B2017.2.DSc/T86 at the Supreme Attestation Commission of the Cabinet of Ministers of the Republic of Uzbekistan.

The dissertation has been prepared at Scientific and Innovation Center of Information and Communication Technologies at the Tashkent University of Information Technologies.

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

Scientific adviser:

Fazilov Shavkat Xayrullaevich

Doctor of technical sciences, professor

Official opponents:

Kabulov Anvar Vasilovich

Doctor of technical sciences, professor

Rustamov Nasim Tulagenovich

Doctor of technical sciences, professor (Kazakhstan)

Nurmuxamedov Tolaniddin Ramziddinovich

Doctor of technical sciences

Leading organization:

Tashkent state technical university

The defense will take place “___” _____ 2018 at _____ on the meeting of Scientific council No. DSc.27.06.2017.T.07.01 at Tashkent University of Information Technologies (Address: 100202, Tashkent city, Amir Temur street, 108. Tel.: (+99871) 238-64-43, fax: (+99871) 238-65-52, e-mail: tuit@tuit.uz).

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

Abstract of dissertation sent out on “___” _____ 2018 y.
(mailing report No. ___ on “___” _____ 2018 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

M.M.Kamilov

Chairman of the academic seminar under the
scientific council awarding scientific degrees,
doctor of technical sciences, academic

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

The aim of the work is the development of methods and algorithms for the formation of informative descriptions of objects using heuristic criteria of informative value of both private and generalized species.

The object of the research are attributes (properties, parameters, characteristics) of objects (processes, phenomena, events) whose values are represented in the form of "object-property" tables.

The novelty of research is as follows:

developed heuristic criteria of informative and non-informative characteristics of the signs and estimated their effectiveness;

defined the properties of the Fisher information criteria used to create methods for determining informative and non-informative sets of characteristics;

developed methods and algorithms for determining informative and non-informative sets of characteristics using a single informational criterion based on the compactness hypothesis;

developed methods and algorithms for determining informative and non-informative sets of characteristics using generalized homogeneous criteria of zero-level and k-th order informativeness;

improved methods based on information content criteria of mixed type and developed algorithms for determining informative and non-informative sets of characteristics;

developed a technique for dividing the initial set of characteristics into groups of informative, low-informative and non-informative features using heuristic criteria.

Implementation of the research results. Based on the software created on the basis of the existing ones, as well as the methods and algorithms of preliminary data processing proposed in the work:

developed software package that allows for pre-processing and storage of data obtained from vibration sensors embedded in the Central Mining Group Navoi Mining and Metallurgical Combine (certificate of the Ministry for Information Technology and Communications №33-8/8663 of 21 December 2017). Using the results of the scientific research, idle time of the KNK-270 conveyor system was reduced by 10% due to prompt notification of dangerous critical situations arising during the operation of this complex;

developed algorithms to improve the quality of speech signals on the basis of pre-processing and filtering implemented in the Investigation Department of the Ministry of Internal Affairs of the Republic of Uzbekistan (the information of the Ministry for Information Technology and Communications №33-8/8663 of 21 December 2017). Based on the use of scientific research results, the processing time of speech signals when translated into textual information was reduced by 20%;

developed software for data mining in the electronic document management system in place at the Ministry of Higher and Secondary Special Education of the Republic of Uzbekistan in the framework of a system of "electronic ministry"

(certificate of the Ministry for Information Technology and Communications №33-8/8663 of 21 December 2017). Based on the use of the results of scientific research, the efficiency of the process of processing data on the teaching staff of higher educational seizures and the formation of relevant proposals through the module of the personnel of the Electronic Ministry system increased twofold.

developed software tools for the formation and preliminary processing of data for the purpose of determining the main indicators of manufactured products are introduced in LLC INTERVETFARM (reference of the Ministry for the Development of Information Technologies and Communications No.33-8/8663 of December 21, 2017). Using the results of the scientific research, the quality of the products was ensured by determining the raw material parameters that must be constantly monitored, which made it possible to reduce by 20% the time required for quality control.

software tools for preprocessing, smoothing and filtering data that ensure the accuracy of contour lines by eliminating various emissions are introduced in LLC Inter sharm plast (reference No.33-8/8663 of the Ministry of Information Technologies and Communications Development of December 21, 2017). Based on the results of scientific research, the time of production based on the use of a specially coded file on the machine, reflecting the main (informative) points allocated on the contour lines was reduced by 15%, which also resulted in a 20% savings in raw materials. production efficiency increased by an average of 1.2 times.

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 175 pages.

ЭЪЛОН ҚИЛИНГАН ИШЛАР РЎЙХАТИ
СПИСОК ОПУБЛИКОВАННЫХ РАБОТ
LIST OF PUBLISHED WORKS

I бўлим (Часть I; Part I)

1. Ш.Х.Фазылов, А.Х.Нишанов, Н.С.Маматов. Методы и алгоритмы выбора информативных признаков на основе эвристических критериев информативности // Монография, Т.: “Fan va texnologiya”, Ташкент, 2017, 132 стр.

2. Комилов М.М., Н.С.Маматов, Худайбердиев М.Х. Методы формирования информативных признаков объектов распознавания на основе критериев Фишеровского типа // Химическая технология. Контроль и управление. – Ташкент, 2010. – № 2. – С. 64-72. (05.00.00; №12)

3. Маматов Н.С., Махкамов А.А. Кулоқ чаноғи тасвирининг информатив белгилари асосида шахсни таниб олиш усул ва алгоритмлари // “Информатика ва энергетика муаммолари” Ўзбекистон журнали, Тошкент, 2011 йил, 6-сон, 34-38 бетлар. (05.00.00; №5)

4. Ravshanov N, Mamatov N., Kurbonov N, Akhmedov D. Parallel computing algorithm for solving the problem of mass transfer in porous medium // European Applied Sciences, Stuttgart (Germany), 2013, №3, pp.40-42. (05.00.00; №2).

5. Н.С.Маматов, Болтибаев Ш.К., Ахмедов Д.Д., Ниёзматова Н.А. Технология параллельного программирования: краткий обзор MPJ Express // Узбекский журнал проблемы информатики и энергетике, –Ташкент. 2013. №3-4. С. 92-97. (05.00.00; №5)

6. Фозилов Ш.Х., Маматов Н.С. L-информатив белгиларни таниб олиш усули // Информатика ва энергетика муаммолари” Ўзбекистон журнали, Тошкент, 2014 йил, 1-2 сон, 10-16 бетлар (05.00.00; №5)

7. Комилов М.М, Худойбердиев М.Х., Маматов Н.С. Взаимосвязь алгоритмов частичной прецедентности // Узбекский журнал проблемы информатики и энергетике, –Ташкент. 2014. №3-4. С. 12-16. (05.00.00; №5)

8. Маматов Н.С., Асраев А.М. Информатив белгилар мажмуасини шакллантириш // “Информатика ва энергетика муаммолари” Ўзбекистон журнали, Тошкент, 2015 йил, 1-2 сон, 68-72 бетлар (05.00.00; №5)

9. Fazilov Sh.X, Mamatov N.S. Selection features using heuristic criteria // International scientific and technical journal “Chemical technology. Control and management №5” and “Journal of Korea multimedia society” South Korea, Seoul – Uzbekistan, Tashkent. 2016. –С.10-15. (05.00.00; №12)

10. Маматов Н.С. Выбор информативных признаков в задаче распознавания образов с использованием модифицированного варианта метода “Дельта” // Узбекский журнал проблемы информатики и энергетике, –Ташкент. 2016. №6. С. 19-22. (05.00.00; №5)

11. Фазылов Ш.Х., Маматов Н.С. Методы и алгоритмы селекция признаков с помощью критериев заданных в виде функционалов k -порядка // Проблемы вычислительной и прикладной математики. –Ташкент. 2017. №2(2). С. 98-102. (05.00.00; №23)

12. Маматов Н.С. Эвристические критерии информативности признаков // Проблемы вычислительной и прикладной математики. –Ташкент. 2017. №4(10). С. 104-109. (05.00.00; №23)

13. Фазылов Ш.Х., Маматов Н.С. Особенности критериев информативности Фишера // Проблемы вычислительной и прикладной математики. –Ташкент. 2017. №4(10). С. 110-118. (05.00.00; №23)

II бўлим (Часть II; Part II)

14. Камиллов М.М., Худойбердиев М.Х., Маматов Н.С. Формирование информативных признаков объектов распознавания на основе однородного критерия Фишеровского типа k -го порядка // Труды девятого международного симпозиума «Интеллектуальные системы», Москва, 2010 г.

15. Фозилов Ш.Х., Маматов Н.С., Ниёзматова Н.А. Информатив белгилар фазосини шакллантиришнинг эвристик мезонлари ва усуллари // Инновацион ғоялар, технологиялар ва лойиҳаларни амалиётга тадбиқ этиш муаммолари, Республика илмий-техник конференцияси, Жиззах, 2014, 417-420 бетлар.

16. Маматов Н.С., Ниёзматова Н.А. Маълумотларга ишлов бериш босқичлари // Инновацион ғоялар, технологиялар ва лойиҳаларни амалиётга тадбиқ этиш муаммолари, Республика илмий-техник конференцияси, Жиззах, 2014, 458-459 бетлар.

17. Fozilov Sh.X., Mamatov N.S. Developing methods and algorithms for forming of informative features' space on base k -type uniform criteria // Eighth World Conference "Intelligent Systems for Industrial Automation", WCIS-2014, 25-27 November 2014, Tashkent, Uzbekistan, pp. 257-261.

18. Hudayberdiyev M.X., Mamatov N.S. Interrelation of parameters estimation algorithms // Eighth World Conference "Intelligent Systems for Industrial Automation", WCIS-2014, 25-27 November 2014, Tashkent, Uzbekistan, pp. 400-402.

19. Фозилов Ш.Х., Маматов Н.С. Маълумотларга дастлабки ишлов бериш масалалари // «Современное состояние и перспективы применения информационных технологий в управлении». –Ташкент. 7-8 сентября 2015. – С. 288-306.

20. Султанов К.С., Маматов Н.С., Ниёзматова Н.А. Программный комплекс диагностики и мониторинга // Современное состояние и перспективы применения информационных технологий в управлении: Доклады Республиканской научно-технической конференции, Джизак, 5-6 сентября 2016 г., с.315-319.

21. Маматов Н.С. Выбор информативных признаков в задаче распознавания образов с использованием модифицированного варианта метода "Дельта" // Современное состояние и перспективы применения информационных технологий в управлении: Доклады Республиканской научно-технической конференции, Джизак, 5-6 сентября 2016 г., с.422-425.

22. Маматов Н.С., Абдурахмонов Б.А. Классификация состояний сложных, многомерных и саморегулярных систем // Современное состояние и перспективы применения информационных технологий в управлении:

Доклады Республиканской научно-технической конференции, Джизак, 5-6 сентября 2016 г., с.425-427.

23. Фозилов Ш.Х., Маматов Н.С. Методы выбора неинформативных признаков // «Алгебра, амалий математика ва ахборот технологиялари масалалари», Республика илмий конференция материаллари, Наманган, 20-21-декабрь, 2016 йил, с.269-273.

24. Абдурахмонов Б.А., Абдуллаева Б.М. Об одном методе формирования пространства неинформативных признаков с помощью эвристических критериев // «Алгебра, амалий математика ва ахборот технологиялари масалалари», Республика илмий конференция материаллари, Наманган, 20-21-декабрь, 2016 йил, с.194-195.

25. Маматов Н.С. Выбор информативных признаков в задаче распознавания образов с использованием модифицированного варианта метода “Дельта” // «Алгебра, амалий математика ва ахборот технологиялари масалалари», Республика илмий конференция материаллари, Наманган, 20-21-декабрь, 2016 йил, с.191-193.

26. Fozilov N.S., Mamatov N.S. Selection features using heuristic criteria // Ninth World Conference “Intelligent Systems for Industrial Automation”, WCIS-2016, 25-27 October 2016, Tashkent, Uzbekistan, pp. 294-298.

27. Mamatov N.S. Working out the methods and algorithms for forming of informative features space with the help of positive - types criteria // International Journal of Mathematical and Computer Modeling, 2016, Vol.21, Issue.1 1222 , pp. 1122-1123.

28. Mamatov N.S. Selection of informative features using heuristic criteria of fisher type // International Journal of Management, Information Technology and Engineering (BEST: IJMITE), 2016, Vol. 4, Issue 11, pp. 7-20.

29. Маматов Н.С. Эвристические критерии информативности признаков // Материалы XVII Международной научно-методической конференции «Информатика: проблемы, методология, технологии». –Воронеж. 9-10 февраля 2017. Том-3, С.114-120.

30. Фозилов Ш.Х., Маматов Н.С. Формализация проблемы снижения размерности признакового пространства в задачах анализа данных // Материалы XVII Международной научно-методической конференции «Информатика: проблемы, методология, технологии». –Воронеж. 9-10 февраля 2017. Том-3, С.173-179.

31. Фозилов Ш.Х., Маматов Н.С. Формализация проблемы снижения размерности признакового пространства // Материалы Республиканской научно-технической конференции «Значение информационно-коммуникационных технологий в инновационном развитии реальных отраслей экономики». -Ташкент. 6-7 апреля 2017. С.382-384.

32. Маматов Н.С. Эвристические критерии информативности // Материалы Республиканской научно-технической конференции «Значение информационно-коммуникационных технологий в инновационном развитии реальных отраслей экономики». -Ташкент. 6-7 апреля 2017. С.319-321.

33. Маматов Н.С., Маҳкамов А.А Шахсни биометрик технологиялар асосида таниб олишнинг асосий йўналишлари // «Современное состояние и

перспективы применения информационных технологий в управлении». – Ташкент. 5-6 сентября 2017. –С. 286-292.

34. Маматов Н.С., Абдурахмонов Б.А. Об одном методе формирования набора малоинформативных признаков // «Современное состояние и перспективы применения информационных технологий в управлении». – Ташкент. 5-6 сентября 2017. –С. 292-295.

35. Маматов Н.С. Маълумотларга дастлабки ишлов бериш масалалари // «Современное состояние и перспективы применения информационных технологий в управлении». –Ташкент. 5-6 сентября 2017. –С. 279-286.

36. Маматов Н.С. Эвристические критерии малоинформативности // «Современное состояние и перспективы применения информационных технологий в управлении». –Ташкент. 5-6 сентября 2017. –С. 295-303.

37. Фозилов Ш.Х., Маматов Н.С. Исследование критериев информативности Фишера // «Современное состояние и перспективы применения информационных технологий в управлении». –Ташкент. 5-6 сентября 2017. –С. 429-441.

38. Мирзаев Н.М., Салиев Э.А., Маматов Н.С. Задачи распознавания объектов, заданных в признаковом пространстве большой размерности // Материалы Республиканской научно-технической конференции «Ахборот коммуникация технологиялари ва сонли моделлаштиришнинг амалий масалалари». –Самарканд. 8-9 сентября 2017. С.244-248.

39. Маматов Н.С. Селекция признаков с помощью критериев Фишерского типа k -го порядка // Материалы Республиканской научно-технической конференции «Ахборот коммуникация технологиялари ва сонли моделлаштиришнинг амалий масалалари». –Самарканд. 8-9 сентября 2017. С.241-244.

40. Абдурахмонов Б.А., Маматов Н.С. Мураккаб объект, ҳодиса ва жараёнлар бошқариш бир дастурий мажмуаси ҳақида // “Фармация: фан, таълим, инновация ва ишлаб чиқариш”, Республика илмий-амалий анжумани, Тошкент, 2017 йил, 16-17-ноябрь.

41. Маматов Н.С. Выбор информативных признаков в задачах распознавания образов // Труды Международной научно-технической конференции «Актуальные проблемы оптимизации и автоматизации технологических процессов и производств», Карши, 2017 год, 17-18-ноябрь.

42. Фозилов Ш.Х., Маматов Н.С. Объект, ҳодиса ва жараёнлар асосий кўрсаткичларини аниқлаш, ҳолатини баҳолаш ва башоратлаш дастурий воситаси. Ўзбекистон Республикаси Интеллектуал мулк агентлиги, Электрон ҳисоблаш машиналари учун яратилган дастурнинг расмий рўйхатдан ўтказилганлиги тўғрисидаги гувоҳнома №DGU 04368.

43. Фозилов Ш.Х., Султанов К.С., Маматов Н.С., Раджабов С.С. Сейсмик объект, ҳодиса ва жараёнлар мониторинги, уларнинг ҳолатини баҳолаш дастурий воситаси. Ўзбекистон Республикаси Интеллектуал мулк агентлиги, Электрон ҳисоблаш машиналари учун яратилган дастурнинг расмий рўйхатдан ўтказилганлиги тўғрисидаги гувоҳнома №DGU 04446.

Автореферат "Ҳисоблаш ва амалий математика муаммолари" илмий журнали тахририятида тахрирдан ўтказилди ҳамда ўзбек, рус ва инглиз тилларидаги матнларини мослиги текширилди.

Бичими: 84x60 ¹/₁₆. «Times New Roman» гарнитура рақамли босма усулда босилди.
Шартли босма табағи: 4. Адади 100. Буюртма № 5.

«ЎзР Фанлар академияси Асосий кутубхонаси» босмахонасида чоп этилди.
100170, Тошкент, Зиёлилар кўчаси, 13-уй.