

TOSHKENT AXBOROT TEXNOLOGIYALARI UNIVERSITETI
HUZURIDAGI ILMIY DARAJALAR BERUVCHI
DSc.09/2025.27.12.T.01.03 RAQAMLI ILMIY KENGASH

DENOV TADBIRKORLIK VA PEDAGOGIKA INSTITUTI

XAYDAROV SHERALI ISLOM O‘G‘LI

SUT BEZI SARATON KASALLIKLARINI ERTA ANIQLASHDA
XATOLIKLARNI MINIMALLASHTIRISHGA ASOSLANGAN
TASHXISLASH ALGORITMI

05.01.03 - Informatikaning nazariy asoslari

TEXNIKA FANLARI BO‘YICHA FALSAFA DOKTORI (PhD)
DISSERTATSIYASI AVTOREFERATI

Toshkent – 2026

**Texnika fanlari bo‘yicha falsafa doktori (PhD) dissertatsiyasi avtoreferati
mundarijasi**

**Оглавление автореферата диссертации доктора философии (PhD) по
техническим наукам**

**Contents of dissertation abstract of doctor of philosophy (PhD) on
technical sciences**

Xaydarov Sherali Islom o‘g‘li

Sut bezi saraton kasalliklarini erta aniqlashda xatoliklarni
minimallashtirishga asoslangan tashxislash algoritmi..... 5

Хайдаров Шерали Ислам огли

Алгоритм диагностики, направленный на минимизацию ошибок
при раннем выявлении рака молочной железы..... 21

Khaydarov Sherali

A diagnostic algorithm aimed at minimizing errors in the early
detection of breast cancer 41

E‘lon qilingan ishlar ro‘uxati

Список опубликованных работ

List of published works..... 45

TOSHKENT AXBOROT TEXNOLOGIYALARI UNIVERSITETI
HUZURIDAGI ILMIY DARAJALAR BERUVCHI
DSc.09/2025.27.12.T.01.03 RAQAMLI ILMIY KENGASH

DENOV TADBIRKORLIK VA PEDAGOGIKA INSTITUTI

XAYDAROV SHERALI ISLOM O‘G‘LI

SUT BEZI SARATON KASALLIKLARINI ERTA ANIQLASHDA
XATOLIKLARNI MINIMALLASHTIRISHGA ASOSLANGAN
TASHXISLASH ALGORITMI

05.01.03 - Informatikaning nazariy asoslari

TEXNIKA FANLARI BO‘YICHA FALSAFA DOKTORI (PhD)
DISSERTATSIYASI AVTOREFERATI

Toshkent – 2026

Texnika fanlari bo'yicha falsafa doktori (PhD) dissertatsiyasi mavzusi O'zbekiston Respublikasi Oliy ta'lim, fan va innovatsiyalar vazirligi huzuridagi Oliy attestatsiya komissiyasida B2025.2.PHD/T5627 raqam bilan ro'yxatga olingan.

Dissertatsiya Denov tadbirkorlik va pedagogika institutida bajarilgan.
Dissertatsiya avtoreferati uch tilda (o'zbek, rus, ingliz (rezyume)) Ilmiy kengash veb-sahifasida (www.tuit.uz) va "Ziyonet" axborot ta'lim portalida (www.ziyonet.uz) joylashtirilgan.

Ilmiy rahbar:

Mamajanov Rahmatilla Yakubjanovich
texnika fanlari nomzodi, dotsent

Rasmiy opponentlar:

Xudayberdiyev Mirzaakbar Xakkulmirzayevich
texnika fanlari doktori, katta ilmiy xodim

Boboyev Lochinbek Boymurotovich
texnika fanlari bo'yicha falsafa doktori (PhD)

Yetakchi tashkilot:

Farg'ona davlat texnika universiteti

Dissertatsiya himoyasi Toshkent axborot texnologiyalari universiteti huzuridagi DSc.09/2025.27.12.T.01.03 raqamli Ilmiy kengashning 2026-yil "10" *iyun* soat 16⁰⁰ dagi majlisida bo'lib o'tadi. (Manzil: 100084, Toshkent shahri, Amir Temur ko'chasi, 108-uy. Tel.: (99871) 207-59-42, e-mail: ilmiy_kengash@tuit.uz).

Dissertatsiya bilan Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universitetining Axborot-resurs markazida tanishish mumkin (416 raqam bilan ro'yxatga olingan). (Manzil: 100084, Toshkent shahri, Amir Temur ko'chasi, 108-uy. Tel.: (99871) 238-64-15).

Dissertatsiya avtoreferati 2026-yil "22" *may* kuni tarqatildi.
(2026-yil "22" - *may* dagi 23 -raqamli reestr bayonnomasi.)



[Handwritten signature]

[Handwritten signature]

M.M.Kamilov

Ilmiy darajalar beruvchi ilmiy kengash raisi, texnika fanlari doktori, professor, O'zFA akademigi

N.A.Egamberdiyev

Ilmiy darajalar beruvchi ilmiy kengash ilmiy kotibi, texnika fanlari bo'yicha falsafa doktori

N.O.Raximov

Ilmiy darajalar beruvchi ilmiy kengash qoshidagi ilmiy seminar raisi, texnika fanlari doktori, dotsent

KIRISH (falsafa doktori (PhD) dissertatsiya annotatsiyasi)

Dissertatsiya mavzusining dolzarbligi va zarurati. Jahonda turli manbalar asosida shakllanayotgan, hajmi katta va murakkab tuzilishga ega bo'lgan ma'lumotlar ustida intellektual tahlilni amalga oshirish, ular tarkibidan muhim va informativ belgilarni ajratib olish hamda tashxislash aniqligini oshirish masalalari ma'lumotlarni intellektual tahlili nazariyasi va amaliyoti rivojida yetakchi o'rinlardan birini egallamoqda. Dunyo miqyosida turli sohalarda qo'llanilayotgan ma'lumotlar to'plamlarida ortiqcha, takroriy va shovqinli belgilarning mavjudligi sinflashtirish natijalarining ishonchliligiga salbiy ta'sir ko'rsatib, tashxislash jarayonida xatoliklar kelib chiqishiga sabab bo'lmoqda. Mazkur holat ma'lumotlarga dastlabki ishlov berish, belgilarning muhimlik darajasini baholash, xato qarorlar sonini kamaytirishga qaratilgan samarali algoritmik yondashuvlarni ishlab chiqish hamda ularni amaliyotga joriy etishni taqozo etadi. Shu jihatdan belgilar fazosining o'lchamini qisqartirish, ortiqcha atributlarni bartaraf etish, eng muhim va informativ belgilar majmuasini shakllantirishga qaratilgan ilmiy izlanishlar ustuvor ahamiyat kasb etadi. Shu bilan birga, obyektlarning xususiyatlarini ifodalovchi parametrlar o'rtasidagi bog'liqliklarni aniqlash, ularning ta'sir darajasini baholash hamda tasniflash aniqligini oshirishga xizmat qiluvchi nazariy yondashuvlar va algoritmik ta'minotdan foydalanish muhim ahamiyatga ega hisoblanadi.

Jahonda ko'p o'lchovli, turli manbalardan shakllanuvchi va murakkab tuzilishga ega ma'lumotlar asosida obyektlarni o'zaro yaqinligi va sinf ichidagi o'xshashlik darajalarini baholash, shuningdek muhim va informativ belgilarni ajratib olishga yo'naltirilgan ilmiy-tadqiqot ishlari olib borilmoqda. Bu borada ortiqcha, shovqinli va takroriy belgilarning mavjudligi natijalarning aniqligi hamda ishonchliligiga salbiy ta'sir ko'rsatishi sababli, belgilarni tanlashning yangi intellektual usullarini ishlab chiqishga ehtiyoj ortib bormoqda. Xususan, minimal qamrovli daraxtga asoslangan obyektlar o'rtasidagi o'xshashlik darajalarini aniqlashda mashinali o'qitish usullari yordamida muhim belgilarni saralash hamda taxshlashdagi xatolik ko'rsatkichlarini kamaytirishga qaratilgan algoritmlarni yaratishga alohida e'tibor berilmoqda.

Respublikamizda ham ma'lumotlarga dastlabki ishlov berish, intellektual tahlil qilish, obyektlarni avtomatik guruhlash va tanib olishga qaratilgan algoritmlar hamda apparat-dasturiy komplekslarni yaratish va amaliyotga joriy etish borasida izchil islohotlar amalga oshirilib, muayyan ilmiy-amaliy natijalarga erishilmoqda. "O'zbekiston - 2030" strategiyasining ustuvor yo'nalishlaridan biri bo'lgan "Iqtisodiyotni raqamlashtirish va zamonaviy texnologiyalarni joriy etish" sohasida, jumladan "...iqtisodiyot tarmoqlari, ijtimoiy soha va davlat boshqaruvi tizimini to'liq raqamlashtirish, IT-xizmatlar eksporti hajmini oshirish hamda raqamli infratuzilmani rivojlantirish..."¹ ustuvor vazifalardan biri sifatida belgilangan. Ushbu vazifalarning amaliy yechimini ta'minlashda, jumladan, tibbiyot, ta'lim, sanoat va qishloq xo'jaligi sohaslarida shakllanuvchi katta hajmli ma'lumotlar

¹ O'zbekiston Respublikasi Prezidentining PF-158-son Farmoni

asosida intellektual qaror qabul qilish tizimlarini ishlab chiqish, tasxislash jarayonlarida xatoliklarni minimallashtirishga asoslangan muhim belgilarni tanlash, tasniflash aniqligini oshirish va xatolik ko'rsatkichlarini kamaytirishga qaratilgan mashinali o'qitish algoritmlarini yaratish muhim ahamiyat kasb etmoqda.

O'zbekiston Respublikasi Prezidentining 2022-yil 28-yanvarda "2022-2026-yillarga mo'ljallangan yangi O'zbekistonning taraqqiyot strategiyasi to'g'risida"gi PF-60-sonli, 2020-yil 29-oktabrdagi "Ilm-fanni 2030-yilgacha rivojlantirish konsepsiyasini tasdiqlash to'g'risida"gi PF-6097-sonli, 2024-yil 14-oktabrdagi PQ-358-sonli "Sun'iy intellekt texnologiyalarini 2030-yilga qadar rivojlantirish strategiyasini tasdiqlash to'g'risida"gi, O'zbekiston Respublikasi Vazirlar Mahkamasining 2025-yil 29-apreldagi 280-sonli "Respublika ixtisoslashtirilgan endokrinologiya ilmiy-amaliy tibbiyot markazi hamda bolalar onkologiyasi, gematologiyasi va immunologiyasi xizmatlarini yanada rivojlantirish chora-tadbirlari to'g'risida"gi, 2023-yil 28-dekabrdagi "Sog'liqni saqlash tizimini raqamlashtirishni jadallashtirish hamda ilg'or raqamli texnologiyalarni joriy etish bo'yicha qo'shimcha chora-tadbirlar to'g'risida"gi PQ-415-sonli qarorlari hamda mazkur faoliyatga tegishli boshqa me'yoriy-huquqiy hujjatlarda belgilangan vazifalarni amalga oshirishga ushbu dissertatsiya ishi muayyan darajada xizmat qiladi.

Tadqiqotning respublika fan va texnologiyalari rivojlanishining ustuvor yo'nalishlariga mosligi. Mazkur tadqiqot Respublika fan va texnologiyalari rivojlanishining IV. "Axborotlashtirish va axborot-kommunikatsiya texnologiyalarini rivojlantirish" ustuvor yo'nalishi doirasida bajarilgan.

Muammoning o'rganilganlik darajasi. So'nggi yillarda sut bezi saraton kasalligini erta aniqlash va tashxislash sohasida ilmiy-amaliy tadqiqotlar jadal rivojlanib, kasallikni dastlabki bosqichlarda aniqlash hamda tashxislash aniqligini oshirishga qaratilgan turli yondashuvlar ishlab chiqilmoqda. Ushbu yo'nalishda kasallik belgilari majmuasini shakllantirish, o'quv tanlanmalarini hosil qilish, tibbiy ma'lumotlar asosida bemor holatini baholash, shuningdek tashxislashga mo'ljallangan model, usul va algoritmlarni yaratish hamda takomillashtirishga alohida e'tibor qaratilmoqda.

Sut bezi saratonini erta aniqlash, diagnostika modellarini ishlab chiqish hamda klinik ma'lumotlar asosida tahlil qilish yo'nalishlarida xorijiy olimlardan R. Barzilay, C.Lehman, K.Lang, S.McKinney, G.Siras, M.Frenk, Li Syuan, Sh.Xuseyn, J.Smit, A.Lopez, K.Vong va olimlar tomonidan tadqiq etilgan. Respublikamizda ma'lumotlarni tahlil qilishning nazariy asoslarini rivojlantirishga M.M.Kamilov, Sh.X.Fozilov, F.T.Adilova, A.X.Nishanov, N.S.Mamatov, B.B.Akbaraliyev va boshqa olimlar muhim hissa qo'shgan. Shuningdek, sut bezi saratonini tashxislash va bashorat qilishga qaratilgan algoritmlar hamda dasturiy vositalarni ishlab chiqish yo'nalishida D.T.Muxamediyeva, O.B.Ro'zibayev va boshqalar tomonidan tadqiqot ishlari amalga oshirilgan.

Mazkur yo'nalishdagi ilmiy ishlar tahlili shuni ko'rsatadiki, mavjud tadqiqotlarda asosan tashxislash modellarini qurishga e'tibor qaratilgan bo'lib, boshlang'ich belgilar fazosidan sinflarni aniq ajratuvchi, yuqori informativlikka ega belgilar majmuasini shakllantirish, ortiqcha va kam ahamiyatli belgilarni chiqarib

tashlash hamda obyektlar o'rtasidagi o'xshashlik darajalarini hisobga olgan holda tashxislash xatoliklarini minimallashtirishga qaratilgan algoritmik yondashuvlar yetarli darajada o'rganilmagan.

Tadqiqotning dissertatsiya bajarilgan oliy ta'lim yoki ilmiy-tadqiqot muassasasining ilmiy-tadqiqot ishlari rejalari bilan bog'liqligi. Dissertatsiya tadqiqoti № SICA-UZ-ISG-2022-004 "Taskin" Surxondaryo viloyati Denov tumani xotin-qizlarni ijtimoiy, iqtisodiy, huquqiy, manfaatlarini himoya qilish axborot markazi Amerika qo'shma shtatlari xalqaro taraqqiyot bo'yicha (USAID) tomonidan moliyalashtiriladigan Markaziy Osiyoda ijtimoiy innovatsiyalar dasturi doirasidagi "Denov404" murojaatlar platformasi ishlab chiqilgan.

Tadqiqotning maqsadi sut bezi saraton kasalliklarini erta aniqlashda tashxislash xatoliklarini minimallashtirish algoritmini ishlab chiqishdan iborat.

Tadqiqotning vazifalari:

sut bezi saratoni kasalliklari klinik ma'lumotlariga oid dastlabki ishlov berish usul va algoritmlari tahlili;

obyektlarni tashxislashda minimal qamrovli daraxt yondashuviga asoslangan klasterlash algoritmini ishlab chiqish;

obyektlarni sinflashtirish xatolik koeffitsiyenti va ehtimollar nazariyasiga asoslangan informativ belgilar majmuasini tanlash algoritmini ishlab chiqish;

informativ belgilar fazosida sinflar obyektlari o'xshashlik darajalari farqi va tashxislashdagi xatoliklarni minimallashtirishga asoslangan tashxislash algoritmini ishlab chiqish;

tibbiyotda sut bezi saratoni kasalliklarini erta tashxislash dasturini ishlab chiqish va amaliyotga tadbiq qilish.

Tadqiqot obyekti ma'lumotlarni intellektual tahlilida tibbiy ma'lumotlarga dastlabki ishlov berish, klasterlash va obyektlarni tashxislashda xatoliklarni minimallashtirish jarayonlari.

Tadqiqot predmeti obyektlarni tashxislashda ma'lumotlarga dastlabki ishlov berish, mikro klasterlashga asoslangan klasterlash, informativ belgilar majmuasini tanlash va xatoliklarni minimallashtirishga asoslangan tashxislash algoritmlari hamda dasturlardan iborat.

Tadqiqotning usullari. Tadqiqot jarayonida tibbiy ma'lumotlarga dastlabki ishlov berish, etalon aniqligini oshirishga qaratilgan mikro klasterlash, informativ belgilarni tanlashda xatolik koeffisienti va ehtimollik nazariyasi, erta tashxislash usullaridan foydalanilgan.

Tadqiqotning ilmiy yangiligi quyidagilardan iborat:

sut bezi saratoniga oid klinik ma'lumotlarda uchraydigan shovqinli, takroriy, yetishmaydigan va nomuvofiq qiymatlarni aniqlash hamda bartaraf etishga qaratilgan dastlabki ishlov berishning adaptiv tozalash algoritmi ehtimollik nazariyasiga asoslangan holda takomillashtirilgan;

obyektlarni tashxislashda yaqinlik va o'xshashlik darajalarini hisobga olingan holda guruhlash aniqligini oshirishga qaratilgan mikro klasterlashning minimal qamrovli daraxtga asoslangan algoritmi ishlab chiqilgan;

obyektlarni sinflashtirishda xatolik koeffitsiyenti va ehtimollar nazariyasi asosida belgilarning ahamiyat darajasini baholash hamda eng muhim xususiyatlarni

ajratib olishga qaratilgan informativ belgilar majmuasini tanlash algoritmi ishlab chiqilgan;

informativ belgilar fazosida sinflar obyektlarining o'xshashlik darajalari farqini aniqlash, tashxislash jarayonida yuzaga keladigan xatoliklarni minimallashtirish, hamda qaror qabul qilish aniqligini oshirishga qaratilgan tashxislash algoritmi ishlab chiqilgan.

Tadqiqotning amaliy natijalari quyidagilardan iborat:

tibbiy ma'lumotlarga dastlabki ishlov berishning adaptiv tozalash va shovqinlarni bartaraf etishga yo'naltirilgan ko'p bosqichli takomillashtirilgan algoritmi kesimida dastur ishlab chiqilgan;

obyektlarni klasterlashda minimal qamrovli daraxtga asoslangan mikroklasterlash va tasniflash algoritmlari kesimida dasturiy vosita ishlab chiqilgan;

sut bezi saratonini erta tashxislashda xatoliklarni minimallashtirilga qaratilgan informativ belgilarni tanlash algoritmidan iborat dastur ishlab chiqilgan;

Tadqiqot natijalarining ishonchliligi. Taklif etilayotgan xatoliklarni minimallashtirish algoritmlarining ishonchliligi ma'lumotlar sifatini oshirishda adaptiv tozalash usullari va shovqinli belgilarni bartaraf etish algoritmlari, informativ belgilarni tanlash samaradorligini ta'minlashda statistik mezonlar, xatolik koeffitsiyentlari va obyektlarni klasterlashda minimal qamrovli daraxtga asoslangan mikroklasterlash, tasniflash algoritmlari, xatoliklarning kamayishi va natijalar aniqligining oshirish algoritmlari ishlab chiqilganliklari bilan tasdiqlangan. Bunda ko'p bosqichli yondashuv tadqiqotning metodologik jihatdan mustahkamligi va natijalarning ishonchliligini kafolatlaydi.

Tadqiqot natijalarining ilmiy va amaliy ahamiyati. Tadqiqot natijalarining ilmiy ahamiyati sut bezi saratoni klinik ma'lumotlariga dastlabki ishlov berishning adaptiv tozalash algoritmini takomillashtirish, minimal qamrovli daraxtga asoslangan mikroklasterlash algoritmini ishlab chiqish hamda xatolik koeffitsiyenti va ehtimollar nazariyasiga asoslangan holda informativ belgilar majmuasini tanlash usullarini yaratish bilan izohlanadi. Shuningdek, informativ belgilar fazosida obyektlar o'rtasidagi o'xshashlik darajalari farqiga asoslangan tashxislash algoritmining ishlab chiqilishi tasniflash aniqligini oshirish va model murakkabligini kamaytirish imkonini beradi, bu esa ma'lumotlarni tahlil qilish jarayonining samaradorligini oshiradi va ilmiy jihatdan muhim natijalar bilan izohlanadi.

Tadqiqot natijalarining amaliy ahamiyati ishlab chiqilgan algoritmlar asosida sut bezi saratonini erta tashxislashga mo'ljallangan "Smartdiagnosis" dasturiy majmuasining yaratilganligi bilan izohlanadi. Ushbu dasturiy vosita tibbiyot muassasalarida klinik va mammografiya ma'lumotlarini tahlil qilish, tashxislash aniqligini oshirish hamda shifokorlarga qaror qabul qilish jarayonida ko'maklashishga xizmat qiladi, bu esa tashxislash jarayonida xatoliklarni kamaytirish va natijalar ishonchliligini oshirish imkonini beradi va amaliy jihatdan muhim natijalar bilan izohlanadi.

Tadqiqot natijalarining joriy qilinishi. Tadqiqotda ishlab chiqilgan sut bezi saraton kasalliklarini erta aniqlashda xatoliklarni minimallashtirishga asoslangan

tashxislash algoritmi va «Smartdiagnosis» dasturiy majmuasidan foydalanish bo'yicha olingan natijalar asosida:

sut bezi saratoniga oid klinik ma'lumotlarga dastlabki ishlov berishning ehtimollik nazariyasiga asoslangan adaptiv tozalash algoritmi ishlab chiqilgan va "SBSK" dasturi O'zbekiston Respublikasi Surxondaryo viloyati Termiz shahar "Respublika ixtisoslashtirilgan onkologiya va radiologiya ilmiy-amaliy tibbiyot markazi Surxondaryo filiali"da tashxislash jarayonlarini takomillashtirish maqsadida amaliyotga joriy qilingan (O'zbekiston Respublikasi Sog'liqni saqlash vazirligining 2025-yil 5-noyabrdagi 02-28/26482-son ma'lumotnomasi). Natijada sut bezi saratonini tavsiflovchi 32 ta simptom belgidan tashxislash uchun eng muhim bo'lgan 18 ta informativ belgilar majmuasi tanlab olingan hamda ushbu belgilar asosida tashxis qo'yishga ketadigan vaqt 20-25 foizga qisqartirilgan va tashxis ishonchliligi 20 foizga oshirilgan;

obyektlarni yaqinlik va o'xshashlik darajalari orqali tashxislashda minimal qamrovli daraxt asosidagi mikroklasterlash algoritmi ishlab chiqilgan va "Smartdiagnosis" dasturi O'zbekiston Respublikasi Surxondaryo viloyati Termiz shahar "Toshkent davlat tibbiyot universiteti Termiz filiali"da tashxislash jarayonlarini takomillashtirish va qaror qabul qilishni qo'llab-quvvatlash maqsadida amaliyotga joriy qilingan (O'zbekiston Respublikasi Sog'liqni saqlash vazirligining 2025-yil 5-noyabrdagi 02-28/26482-son ma'lumotnomasi). Natijada tanlangan informativ belgilar majmuasidan foydalanib kasallikka tashxis qo'yishga ketadigan vaqt 1,5 barobarga qisqartirilgan va tashxislash samaradorligi 1,7 barobarga oshirilgan.

obyektlarni tashxislashdagi xatoliklarni minimallashtirishda informativ belgilar fazosida xatolik koeffisienti kamaytirishga qaratilgan tashxislash algoritmi ishlab chiqilgan va "Smartdiagnosis" dasturi O'zbekiston Respublikasi Surxondaryo viloyati Termiz shahar "O'tan Polvon diagnostika va davolash markazi"da tashxislash samaradorligini oshirish maqsadida amaliyotga joriy qilingan (O'zbekiston Respublikasi Sog'liqni saqlash vazirligining 2025-yil 5-noyabrdagi 02-28/26482-son ma'lumotnomasi). Natijada ishlab chiqilgan algoritmlar asosida sut bezi saratonini tashxislash jarayonlari avtomatlashtirildi hamda diagnostika jarayonining ishonchliligi 20 foizga oshirishi tasdiqlangan.

Tadqiqot natijalarining aprobatsiyasi. Dissertatsiyaning asosiy nazariy hamda amaliy natijalari 5 ta xalqaro va 3 ta respublika ilmiy-texnika hamda ilmiy-amaliy anjumanlarida muhokama qilingan.

Tadqiqot natijalarining e'lon qilinganligi. Tadqiqot mavzusi bo'yicha jami 20 ta ilmiy ish chop etilgan, shulardan, O'zbekiston Respublikasi Oliy attestatsiya komissiyasining doktorlik dissertatsiyalari asosiy ilmiy natijalarini chop etishga tavsiya etilgan ilmiy nashrlarda 9 ta maqola, 3 ta xorijiy va 6 ta respublika jurnallarida nashr qilingan hamda 3 ta EHM uchun dasturiy mahsulot qayd etilganligi haqida guvohnoma olingan.

Dissertatsiyaning tuzilishi va hajmi. Dissertatsiya kirish, to'rt bob, xulosa, foydalanilgan adabiyotlar ro'yxati va ilovalardan iborat. Dissertatsiyaning hajmi 118 betni tashkil etadi.

DISSERTATSIYANING ASOSIY MAZMUNI

Kirish qismida dissertasiya mavzusining dolzarbligi va zaruriyati asoslangan, tadqiqotning respublika fan va texnologiyalari rivojlanishining ustuvor yo'nalishlariga mosligi, mavzuning o'rganilganlik darajasi, dissertasiya bajarilgan oliy ta'lim muassasi ilmiy-tadqiqot ishlari rejalari bilan bog'liqligi ko'rsatilgan, dissertasiyaning maqsadi, vazifalari, obyekti, predmeti, usullari tavsiflangan, tadqiqotning ilmiy yangiligi va amaliy natijalari, tadqiqot natijalarining ishonchliligi, olingan natijalarining ilmiy va amaliy ahamiyati ochib berilgan, tadqiqot natijalarining joriy qilinishi, tadqiqot natijalarining aprobatsiyasi, nashr etilgan ishlar va dissertasiya tuzilishi bo'yicha ma'lumotlar keltirilgan.

Dissertatsiyaning **“Sut bezi saraton kasalliklarini erta aniqlash usul va algoritmlari tahlili”** deb nomlangan birinchi bobida zamonaviy texnologiyalarda mashinali o'qitish usullari, saraton kasalligini aniqlashda ancha yuqori aniqlik va samaradorlikni ta'minlashi bayon etilgan. Ushbu texnologiyalar tibbiy ma'lumotlarni tahlil qilish, kasallikning diagnostik xususiyatlarini aniqlash va qarorlarni qabul qilish jarayonlarini avtomatlashtirish imkoniyatini beradi.

1.1-paragrafda sut bezi saraton kasalliklarini simptonlarni mashinali o'qitishga tayyorlash bosqichlari, belgilarda axboriy muhimlik tushunchasining ilmiy asosi, klinik belgilarning informativ muhimligi to'g'risida batafsil tahliliy ma'lumotlar keltirilgan.

1.2-paragrafda saraton kasalliklarini erta aniqlashdagi zamonaviy yondashuvlar, saraton kasalliklarini erta aniqlashning davolash samaradorligini oshirishdagi ahamiyati hamda tasviriy diagnostika, biomarkerlar va genetik ma'lumotlar asosida zamonaviy sun'iy intellekt algoritmlari yordamida yuqori aniqlikdagi tashxis qo'yish imkoniyatlari tahlil qilingan.

1.3-paragrafda saraton kasalliklarini aniqlash mezonlari va algoritmlari, saraton kasalliklarini aniqlashda qo'llaniladigan klinik, laboratoriya, tasviriy, gistologik va biomarker mezonlari, shuningdek sun'iy intellekt va mashinani o'rganish algoritmlariga asoslangan zamonaviy diagnostika yondashuvlari o'rganilgan va asoslab berilgan.

1.4-paragrafda dissertasiya ishining asosiy masalasi qo'yilishi bayon etilgan bo'lib, tashxislashdagi xatolarni minimallashtirishga asoslangan optimizatsiya masalasining qo'yilishi va ularni yechishga doir yondashuvlar keltirilgan.

Faraz qilaylik, X o'quv tanlanma quyidagi ko'rinishida berilgan bo'lsin, ya'ni

$$\begin{aligned} X &= \bigcup_{p=1}^r X_p, X_p \cap X_q = \emptyset, p \neq q, \quad p, q = \overline{1, r}, \\ X_p &= \{x_{pi} = (x_{pi}^1, x_{pi}^2, \dots, x_{pi}^N): i = \overline{1, m_p}\} \end{aligned} \quad (1)$$

bu yerda r - sinflar soni, X_p -sinflar ($p = \overline{1, r}$), m_p - X_p sinfdagi obyektlar soni ifodalaydi. Obyektlarni tavsiflovchi optimal informativ tizimlarni shakllantirishda foydalaniladigan $\lambda = (\lambda^1, \lambda^2, \dots, \lambda^N)$ -informativlik vektoriga tegishli ba'zi tushunchalar keltiriladi. Berilgan λ vektorga nisbatan obyektlarni tavsiflovchi informativ tizimni quyidagicha (2) bilan aniqlanadi.

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

Ta'rif. Agar $\sum_{j=1}^N \lambda^j = \ell$ bo'lsa, u holda $\lambda = (\lambda^1, \lambda^2, \dots, \lambda^N)$ vektor ℓ informativ deyiladi. Barcha ℓ informativ vektorlar to'plamini Λ^ℓ orqali ifodalaydi.

$$\Lambda^\ell = \{ \lambda: \sum_{j=1}^N \lambda^j = \ell, \lambda^j \in \{0,1\}, j = \overline{1, N} \} \quad (3)$$

(3) ifodada ko'rinib turibdiki, Λ^ℓ to'plam quvvati quyidagiga teng $|\Lambda^\ell| = C_N^\ell = \frac{N!}{\ell!(N-\ell)!}$, u holda $\Lambda^\ell: X \rightarrow X|_{\Lambda^\ell} = \{X|_\lambda \mid \lambda \in \Lambda^\ell\}$ ifodalaydi. Bundan tashqari, ℓ informativ vektorlar Λ^ℓ to'plami berilgan bo'lib, ixtiyoriy $\lambda \in \Lambda^\ell$ vektorga mos obyektlarni tavsiflovchi informativ tizimlar quyidagicha aniqlanadi.

$$\lambda: X \rightarrow X|_\lambda, \quad X|_\lambda = \{x \in X: x = (\lambda^1 x^1, \lambda^2 x^2, \dots, \lambda^N x^N), \sum_{i=1}^N \lambda^i = \ell\}. \quad (4)$$

Ma'lumotlar intellektual tahlilida o'quv tanlanma obyektlari turli o'lchovdagi belgilar fazosida, ya'ni bitta o'lchovdagi belgilar fazosidan boshqa o'lchovli belgilar fazosiga o'tganda, o'zlarini qanday tutishini aniqlash muhim hisoblanadi. Masalan, sinflashtirish masalasi hal qilinayotganda, belgilar fazosini o'lchovini o'zgartirish natijasida, ba'zi obyektlar o'z sinflariga emas, balki boshqa sinfga o'tib qolishi mumkin, ya'ni

$$\exists x \in X, \lambda \in \Lambda^{\ell_1}, \mu \in \Lambda^{\ell_2}, \ell_1 \neq \ell_2 \text{ uchun } x|_\lambda \in X_p, x|_\mu \in X_q, p \neq q.$$

Bunday holatlarni aniqlash va ularning oldini olish maqsadida belgilar fazosi o'lchovini kamaytirish jarayonida sinflashtirishdagi xatoliklar inobatga olinadi.

Faraz qilaylik, ℓ informativ vektorni λ vektorga nisbatan sinflashtirishdagi xatoliklar koeffitsiyenti $\theta(\ell)|_\lambda$ va o'z sinfini topa olmagan, ya'ni noto'g'ri sinflashtirilgan obyektlar soni $\kappa(\ell)|_\lambda$ bo'lsin. Sinflashtirishdagi xatoliklar koeffitsiyenti bilan noto'g'ri sinflashtirilgan obyektlar soni orasidagi bog'liqlikni quyidagicha aniqlaymiz.

$$\theta(\ell)|_\lambda = \frac{\kappa(\ell)|_\lambda}{m} \quad (5)$$

bu yerda $m = \sum_{p=1}^r m_p$ - o'quv tanlanmadagi obyektlarning umumiy soni.

Izoh. Agar aynan qaysidir sinf (sinflar) obyektlari boshqa sinfga yoki aksincha, boshqa sinf obyektlari shu sinf (sinflar)ga adashib o'tib qolmasligi muhim bo'lsa, u holda sinflashtirishdagi xatoliklar koeffitsiyentini aniqlashda sinflarning muhimlik darajasini ham e'tiborga olish lozim bo'ladi.

Masalan X_p sinf obyektini boshqa sinfga adashib o'tib qolmaslik muhimlik darajasi α_p va aksincha boshqa sinf obyektining bu sinfga adashib o'tib qolmasligining muhimlik darajasi β_p bo'lsin. U holda X_p sinf uchun noto'g'ri sinflashtirilgan obyektlar sonini muhimlik darajasiga ko'ra quyidagicha aniqlash mumkin.

$$\kappa(\ell)|_\lambda(X_p) = \alpha_p \cdot t_p^{out}|_\lambda + \beta_p \cdot t_p^{in}|_\lambda \quad (6)$$

bu yerda t_p^{out} va t_p^{in} mos ravishda X_p sinfdan xato chiqib ketgan va qo'shib qolgan obyektlar soni ifodalaydi, (5) ga asosan (6) quyidagi ko'rinishga ega bo'ladi:

$$\theta(\ell)|_\lambda = \frac{\sum_{p=1}^r (\alpha_p \cdot t_p^{out}|_\lambda + \beta_p \cdot t_p^{in}|_\lambda)}{m} \quad (7)$$

Nominal belgilar fazosida sinflashtirishdagi xatoliklarni minimallashtirishda informativ belgilar majmuasini tanlash talab etilsin, ya'ni quyidagi optimizatsiya masalasi yechilsin.

$$\begin{cases} \theta(\ell)|_{\lambda} = \frac{x(\ell)|_{\lambda}}{m} \rightarrow \min \\ \Lambda^{\ell} = \{ \lambda: \sum_{j=1}^N \lambda^j = \ell, \lambda^j \in \{0,1\}, j = \overline{1, N} \} \end{cases} \quad (8)$$

Dissertatsiyaning “**Tibbiy obyektlarni sinflashtirishda xatoliklarni minimallashtirishga asoslangan algoritmlar**” deb nomlangan ikkinchi bobida sut bezi saratoni bo‘yicha tibbiy ma’lumotlarni adaptiv tozalash algoritmlari asosida umumiy tanlanma shakllantirildi. Shuningdek, takomillashtirilgan MSTL klasterlash algoritmi asosida mikro-klasterlash va makro-klasterlash jarayonlarini amalga oshirish imkoniyatlari asoslab berilgan.

2.1- paragrafda sut bezi saratoniga oid tibbiy ma’lumotlarni adaptiv tozalash algoritmi mavzusida umumiy tanlanmadagi obyektlarni ichidagi xato kirib qolgan belgilarni tozalab ehtimollikga asoslanib eng ko‘p uchraydigan belini qo‘yish orqali yangi umumiy tanlanmani hosil qilindi. “Sut bezi saratoni kasalligi belgilari va ularning qabul qiladigan qiymatlari” keltirilgan(1-jadval).

1-jadval.

Sut bezi saratoni kasalligi belgilari va ularning qabul qiladigan qiymatlari

Belgilar	Sut bezi onkologiyasi kasallik belgilari nomlanishi	Belgilarning qabul qilishi mumkin bo‘lgan qiymatlari
x^1	Sizda kasallik nechi yoshda boshlangan?	1. Kasallik yuq 2. 17-30 yoshda 3. 30-50 yoshda 4. 50 yoshdan keyin
x^2	Sizni sut bezingizda xosila(o‘sma) bormi?	1. Yo‘q
x^3	Sizda umumiy xolsizlik(darmonsizlik) kuzatiladimi?	1. Yo‘q 2. Yengil(o‘rta)
...		
x^{32}	Siz saqlanish uchun muntazam ravishda kontrasepsiya yoki terapevtik garmon preparatlarini qabul qilasizmi?	1. Yo‘q 2. Ha

Sut bezi saratoniga oid tibbiy ma’lumotlarni adaptiv tozalash algoritmi doirasida klinik belgilar tartibli nominal qiymatlar asosida baholangan bo‘lib, ushbu belgilar klinik tajriba va amaliyotga tayangan holda shifokor-ekspertlar tomonidan oldindan kodlangan. Tadqiqotda shakllantirilgan umumiy o‘quv tanlama real klinik kuzatuvlarga asoslanib, unda har bir satr alohida bemorni, har bir ustun esa muayyan klinik belgini ifodalaydi hamda keyingi intellektual tahlil va klasterlash algoritmlari uchun asosiy ma’lumotlar bazasi vazifasini bajaradi.

Adaptiv tozalash algoritmi quyidagi qadamlarda amalga oshiriladi.

1-qadam. Boshlang‘ich ma’lumotlarni kiritish

$$X = \{x_1, x_2, \dots, x_m\}, x_i = (x_i^1, \dots, x_i^N), i = \overline{1, m}$$

2-qadam. Har bir nominal belgi qiymati mumkin bo‘lgan qiymatlar to‘plamiga mosligi tekshiriladi.

$$x_i^j = \begin{cases} x_i^j, & x_i^j \in \Omega_j, \\ \emptyset, & \text{aks holda} \end{cases} \quad (9)$$

bu yerda Ω_j - j-belgining mumkin bo'lgan qiymatlar to'plami.

3-qadam. Adaptiv tozalash va bo'sh qiymatlarni to'ldirish. Har bir j -belgi uchun eng ko'p uchraydigan qiymat aniqlanadi

$$v_j = \operatorname{argmax}_{v \in \Omega_j} P(x^j = v) \quad (10)$$

bu yerda

$$P(x^j = v) = \frac{\operatorname{count}(x^j=v)}{m} \quad (11)$$

Bo'sh qiymatlar ochiq qoldirilmaydi va quyidagi qoida asosida to'ldiriladi.

$$x_i^j = \begin{cases} x_i^j, & x_i^j \neq \emptyset \\ v_j, & \text{aks holda} \end{cases} \quad (12)$$

4-qadam. Yakuniy tanlanmani shakllantirish. Adaptiv tozalash va o'xshashlik darajasini baholash bosqichlari natijasida umumiy tanlanma shakllantirildi(1-rasm).

$$X = \tilde{X} = \{\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_m\} \quad (13)$$

№	Sizda kasallik nechi yoshda boshlangan?	Sizni sut bezingizda xosila(o'sma) bormi?	Sizda umumiy xosizlik(darmonsizlik) kuzatladimi?	Sizda kasallikni qanchadan beri rivojlangan deb bilasiz?	Sizni nashingizda (onasi-opa-singillar, amma -volalarda) kasallar bo'lganmi?	Sizni yelkaming pastki sohasidagi og'riqlar bormi?	Xosila(o'sma)ning qo'ltiq ostida kattalashgan limfa tugunlari mavjudmi?	Sizning sut bezining o'lchami qanday?	...	Turmushga necha yoshda chiqqansiz?	Siz nechi yoshda farzandli bo'lgansiz?	Sizni vazningizda o'zgarish bormi?	Sizni sut bezingizda tugunli xosilani sezasizmi?	Sut bezidagi xosilani (bez) ushlab kurganizda harakatlanadimi?	Sut bezidagi xosilani chegaralari aniqmi	Sut bezidagi o'smaning holati qanday	Saqlanish uchun siz muntazam ravishda konroseptom garmon preparatlarini qabul qilasizmi
	X1	X2	X3	X4	X5	X6	X7	X8	...	X25	X26	X27	X28	X29	X30	X31	X32
1	4	2	2	2	1	2	2	2	...	2	2	3	2	1	2	1	1
2	3	2	2	2	1	1	1	1	...	3	3	1	1	2	2	1	2
3	3	2	2	3	1	1	2	2	...	1	1	1	1	2	2	1	1
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	...
563	3	2	3	2	1	1	2	1	...	2	2	1	2	1	2	3	1
564	4	2	3	2	1	1	2	2	...	2	2	1	2	2	2	3	1
565	3	2	3	3	1	3	1	1	...	2	2	1	2	2	1	3	1
566	3	2	3	2	1	1	2	2	...	1	1	1	2	1	1	3	1
567	3	2	3	2	1	1	2	1	...	2	2	3	2	1	1	3	1

1-rasm. Umumiy tanlama

Taklif etilgan adaptiv tozalash algoritmi sut bezi saratoniga oid nominal klinik ma'lumotlarni qo'shimcha kodlash amallarini talab etmasdan ularning sifatini yaxshilashga yo'naltirilgan. Ushbu yondashuv ma'lumotlardagi noaniqlik va bo'sh qiymatlarni bartaraf etish orqali yuqori o'lchamli nominal fazoda klasterlash va tasniflash jarayonlarining aniqligi hamda barqarorligini oshirishni ta'minlaydi.

2.2-paragrafda minimum spanning tree (MST) asosidagi mikroagregatsiya algoritmi orqali ma'lumotlar kamida k ta elementdan iborat guruhlariga ajratiladi. Har bir guruh o'z ichidagi elementlar o'rtasidagi o'xshashlik darajasiga asoslangan holda aniqlangan vakil qiymat bilan ifodalanadi. Guruhlarni shakllantirish jarayoni

ma'lumot yo'qotilishini minimallashtirish mezoniga asoslanadi. MST tuzilmasi asosida obyektlar o'rtasidagi eng yaqin bog'lanishlar aniqlanadi.

Takomillashtirilgan MSTL klasterlash algoritmi.

1-qadam. Faraz qilaylik, X o'quv tanlanma quyidagi ko'rinishida berilgan bo'lsin, ya'ni

$$X = \bigcup_{p=1}^r X_p, X_p \cap X_q = \emptyset, p \neq q, p, q = \overline{1, r}, \quad (14)$$

$$X_p = \{x_{pi} = (x_{pi}^1, x_{pi}^2, \dots, x_{pi}^N) : i = \overline{1, m_p}\},$$

bu yerda r - sinflar soni, X_p - sinflar ($p = \overline{1, r}$), m_p - X_p sinfdagi obyektlar soni, x_{pi} - p -sinfdagi i -obyekti, x_{pi}^j - p - sinfdagi i - obyektning j - belgisi ifodalaydi.

Klasterlash masalasini yechish uchun obyektlararo o'xshashlik funksiyasi kiritildi.

2-qadam. Nominal belgilar fazosida x_{pi} va x_{pq} obyektlarining j -belgisiga mos o'xshashligini ko'rsatuvchi kattalik

$$\rho_{pi}^j(x_{pi}, x_{pq}) = \begin{cases} 1, & \text{agar } x_{pi}^j = x_{pq}^j \\ 0, & \text{aks holda} \end{cases} \quad (15)$$

bu yerda $p = \overline{1, r}$; $i \neq q = \overline{1, m_p}$; $j = \overline{1, N}$; Mazkur kattalik har bir belgi bo'yicha obyektlar o'xshashligini ifodalaydi. Obyektlar orasidagi umumiy o'xshashlik ushbu qiymatlarning o'rtacha arifmetikasi sifatida aniqlanadi.

$$\rho(x_{pi}, x_{pq}) = \frac{1}{N} \sum_{j=1}^N \rho_{pi}^j(x_{pi}, x_{pq}) \quad (16)$$

3-qadam. Har bir x_{pi} obyekt uchun r_p o'xshashlik chegarasi ichidagi qo'shnilar aniqlanadi. Radius r_p qiymati X_p sinfdagi obyektlarning k -eng yaqin qo'shnilariga mos keluvchi o'xshashlik qiymatlarining o'rtacha qiymati asosida aniqlanadi.

$$r_p = \frac{1}{m_p} \sum_{i=1}^{m_p} \rho^{(k)}(x_{pi}), \quad K_i = \{x_{pq} \in X \mid \rho_{pi}(x_{pi}, x_{pq}) \geq r_p\} \quad (17)$$

Agar

$$|K_i| \geq \text{MinPts}, \quad (18)$$

shart bajarilsa, x_{pi} obyekt mikro - klaster markazi bo'ladi.

$$|K_i| \geq k \Rightarrow x_{pi} \text{ mikro-klaster markazi}$$

$\rho^{(k)}(x_{pi})$ - x_{pi} obyektning k -eng yaqin qo'shnisiga mos umumiy o'xshashlik qiymati, m_p - X_p sinfidagi obyektlar soni, $|K_i|$ - ushbu markazga mos kelgan qo'shnilar soni ifodalaydi.

4-qadam. Mikroagregatsiya jarayonida klasterlash samaradorligini baholash uchun yo'qotish funksiyalari sifatida SSE(sum of squared errors) va SST(total sum of squares) ko'rsatkichlaridan foydalaniladi. Ushbu ko'rsatkichlar obyektlarning klaster ichidagi va umumiy tarqoqlik darajasini miqdoriy baholash imkonini beradi.

$$\sum_{p=1}^r m_p = M, SSE = \frac{1}{M} \sum_{p=1}^r \sum_{i=1}^{m_p} (\rho_{pi}(x_{pi}, \bar{x}_p))^2, SST = \frac{1}{m_p} \sum_{i=1}^{m_p} \rho_{pi}(x_{pi}, \bar{x})^2 \quad (19)$$

SSE-klaster ichidagi obyektlarning klaster markaziga nisbatan hisoblangan ichki tarqoqlik mezoni.

SST-barcha obyektlarning umumiy markazga nisbatan hisoblangan umumiy

tarqoqlik mezoni.

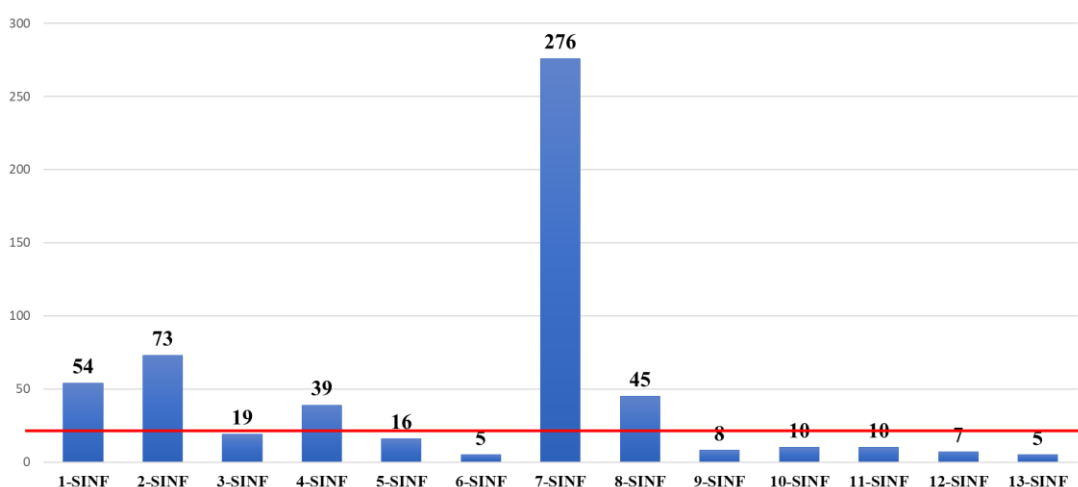
bu yerda x_{pi} -p-sinfning i -obyekti, \bar{x}_p -p-sinfning markaziy (o'rtacha) qiymati, \bar{x} - barcha obyektlarning umumiy o'rtacha qiymati; m_p - p-sinfidagi obyektlar soni ifodalaydi.

5-qadam. L klasterlash sifatini baholovchi ko'rsatkich bo'lib, SSE va SST nisbatini ifodalaydi.

$$L = \frac{SSE}{SST} \rightarrow \max \quad (20)$$

L qiymatining ortib borishi klasterlash aniqligining oshishini bildiradi. $L = 1$ bo'lsa mukammal klasterlash ega bo'ladi.

Ishlab chiqilgan algoritm asosida klasterlash jarayoni amalga oshiriladi. Natijada 2-rasmda sinflardagi obyektlar sonining sinflar bo'yicha taqsimoti diagrammasi hosil bo'ladi.



2-rasm. Sinflardagi obyektlar sonining sinflar bo'yicha taqsimoti

2.3-paragrafda sut bezi saratoni kasalligi o'quv tanlanmasini normallashtirishda sintetik obyektlarni shakllantirish algoritmi va ularning mazmunini aniqlab beruvchi mezonlar, yarim sintetik o'quv tanlanmani yaratish algoritmi ishlab chiqilgan bo'lib, ushbu algoritm nominal belgilar fazosi uchun sintetik obyektlar yaratishda obyektlarning baholash orqali yangi kengaytirilgan datasetni yaratishga qaratilgan.

Sintetik o'quv tanlanma yaratish algoritmi

1-qadam. Umumiy o'quv tanlanma obyektlari ma'lumotlar bazasiga kiritiladi. Boshlang'ich ma'lumotlar bazasi barcha $x_i \in X, i = \overline{1, M}$ obyektlar kesimida shakllantiriladi.

2-qadam. Ma'lumotlarga dastlabki ishlov berish jarayoni amalga oshiriladi. Bunda ma'lumotlarni tozalash ya'ni to'plangan ma'lumotlarda xatolar, yo'qolgan qiymatlar, nusxalar yoki noaniq ma'lumotlar bo'lishi mumkin. Bular to'g'rilanadi va ma'lumotlar normirovka qilinadi.

3-qadam. Umumiy o'quv tanlanma oldindan quyidagi obyektlar $x_{p1}, x_{p2}, \dots, x_{pm_p} \in X_p, p = \overline{1, r}$ sinflarga ajratiladi.

4- qadam. Shakllantirilgan har bir sinf uchun obyektlarning o'rta o'xshashligi $\Gamma_p(X_p), p = \overline{1, r}$ ning qiymatlari aniqlanadi.

5-qadam. Yangi \bar{x} obyektini X_p , $p = \overline{1, r}$ sinflari uchun ushbu sinfdagi obyektning o'rtacha o'xshashlik darajasi $\Gamma_p(X_p)$ ning ko'tarilishini ta'minlagan muhim belgilari saqlangan holda yaratiladi.

6-qadam. Agarda p -sinfda obyektlar soni $m_p = 1$ bo'lsa, u holda sintetik obyekt yaratilganda sinfdagi obyektning boshqa sinflar obyektlaridan ajralib turadigan belgilari saqlangan holda yaratiladi.

7-qadam. 5-6 - qadamlar har bir sinfga H ta yangi sintetik obyektlar qo'shilgunicha davom etadi. Natijada gibrid tanlamadan iborat sinflar paydo bo'ladi.

Dissertatsiya ishining **“Sut bezi saratoni kasalliklarini erta aniqlash algoritmlari va inovatsion yondashuvlar”** deb nomlangan uchinchi bobida tasniflash va bashoratlash tizimlarida informativ belgilar majmuasini tanlash xatoliklarni minimallashtirishda muhim ahamiyatga ega bo'lib, kam ahamiyatli va shovqinli belgilarni chiqarib tashlash orqali modelning aniqligi va umumlashtirish qobiliyatini oshiradi.

3.1-paragrafda tasniflashda informativ belgilar majmuasini tanlashning xatoliklarni minimallashtirishga asoslangan algoritmini takomillashtirish orqali tasniflashdagi xatoliklarni kamaytirish masalasi o'rganiladi. Bunda ma'lumotlar o'lchamini pasaytirish, diagnostik va bashorat qilish aniqligini oshirish hamda hisoblash samaradorligini yaxshilash kabi vazifalarni amalga oshirish ko'zda tutilgan. Shuningdek, amaliy tajribada qo'llanilayotgan algoritmlar tahlil qilinadi va ularning xatolikka ta'siri solishtirildi.

Xatolikni minimallashtirishga qaratilgan informativ belgilarini tanlash algoritmi quyidagicha qadamlarda hisoblanadi.

1-qadam. Boshlang'ich ma'lumotlar.

X - o'quv tanlanma; $X_p(p = \overline{1, r})$ - sinflar; r - sinflar soni; $k, \ell, N, \theta, T, i, p, \mu$.

2-qadam. $\lambda \in \Lambda^N$ uchun $\kappa(\ell)|_\lambda$ ba $\theta = \theta(\ell)|_\lambda$ hisoblanadi.

3-qadam. $i = k$

4-qadam. $v = 1$ da $p_v = (p_v^1, p_v^2, \dots, p_v^N)$ uchun $p_v^j = \frac{1}{N}; j = \overline{1, N}$ qiymatlar qabul qilinadi.

5-qadam. Barcha $\mu_i = \left\{ \underbrace{0, 0, \dots, 0}_{N-i}, \underbrace{1, 1, \dots, 1}_i \right\}$, $\mu \frac{N!}{i!(N-i)!} = \left\{ \underbrace{1, 1, \dots, 1}_i \underbrace{0, 0, \dots, 0}_{N-i} \right\}$, $\mu \in \Lambda^i$ vektorlar uchun $\theta(N-i)|_{\lambda-\mu}$ ba $\theta(i)|_\mu$ lar hisoblanadi.

6-qadam. Barcha μ vektorlar uchun agar $\theta(N-i)|_{\lambda-\mu} \leq \theta$ o'rinli bo'lsa, $p_{v+1}(\mu) = p_v + \frac{(p_v, \mu)}{N-i} * p_v|_{\lambda-\mu} - p_v|_\mu$ deb o'zlashtiriladi.

7-qadam. $p_{v+1} = \frac{1}{T} \sum_{\mu} p_{v+1}(\mu)$ hisoblanadi, bu yerda $T = p_{v+1}(\mu)$ vektorlar soni $p_v = (p_v^1, p_v^2, \dots, p_v^N)$ vektor uchun tekshiriladi va shunga mos ravishda i ta belgi belgilar fazosidan chiqariladi.

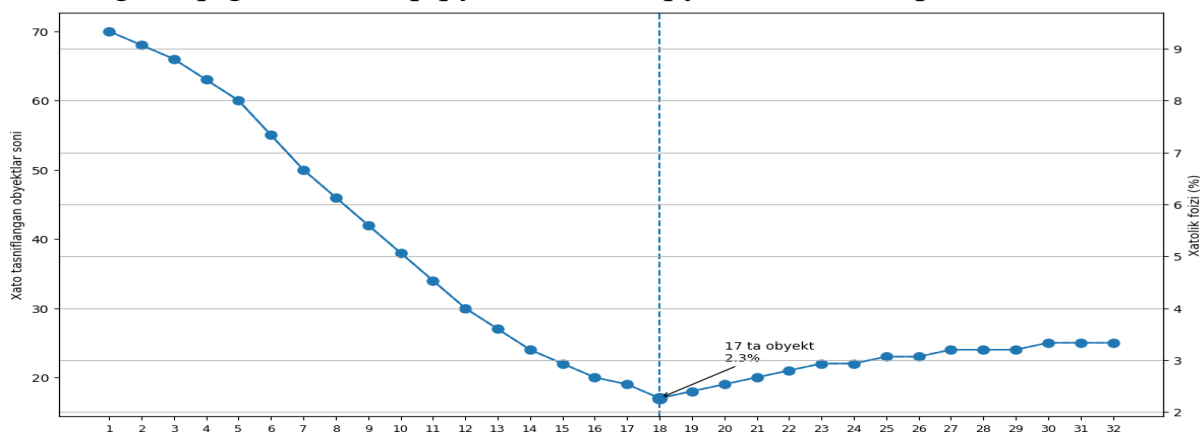
8-qadam. $N = N - i$ ba $\lambda = (1, 1, \dots, 1) \in \Lambda^N$ deb yangi qiymat o'zlashtiriladi.

9-qadam. Agar $i < N - \ell$ bo'lsa, u holda 3-qadamga qaytiladi.

10-qadam. Agar $i > N - \ell$ bo'lsa, u holda $i = N - \ell$ deb olinadi va 4-qadamga qaytiladi.

11-qadam. Chiqish parametri λ optimal vektor.

Belgilar soni 1 tadan 18 taga bosqichma-bosqich oshirilganda tasniflash xatoligi sezilarli darajada kamayib borgan bu holat modelning yangi axborot olishi bilan o'rganish sifati yaxshilanishini ko'rsatadi. 18-belgiga kelib xatolik eng quyi darajaga 2.3% ga (17 ta obyekt) tushgan va bu nuqta optimal konfiguratsiya sifatida belgilangan. 19-belgidan 32-belgigacha bo'lgan oraliqda esa xatolik kamaymay, aksincha, asta-sekin oshib borgan: bu hodisa modelga ortiqcha va shovqinli belgilar kiritilganda yuzaga keladigan haddan ortiq o'rganish effektining klassik ko'rinishi hisoblanadi. Shunday qilib 3-rasmda keltirilgan 32 ta nominal belgidan faqat 18 tasi modelning aniqligi uchun haqiqiy informativ qiymat sifatida topildi.



3-rasm. Tasniflash xatoligining ℓ bog'liq o'zgarishi

3.2-paragrafda xatoliklarni minimallashtirishga qaratilgan tashxislash algoritmi keltirilgan.

Faraz qilaylik N -o'lchovli nominal belgilar fazosida sut bezi saratoni kasalligining o'quv tanlanmasi $x_{p1}, x_{p2}, \dots, x_{pm_p} \in X_p, p = \overline{1, r}$ berilgan bo'lsin. O'quv tanlanmadagi har bir obyekt $x_{pi} = (x_{pi}^1, x_{pi}^2, \dots, x_{pi}^N)$, $i = \overline{1, m_p}$ bemorlar N -o'lchovli nominal belgilar fazosida berilgan. Bu yerda x_{pi} , p -sinfning i -obyekti deb o'qiladi, N - obyektarni tashkil etuvchi belgilar sonini bildirsa, m_p - p sinfdagi obyektlar sonini bildiradi. Demak, $x_{pi} = (x_{pi}^1, x_{pi}^2, \dots, x_{pi}^N)$ bemorlari p -sinfdagi i ($i = \overline{1, m_p}$) - obyektarni tashkil etadi.

Masalaning qo'yilishi. Nazorat tanlanmasi $W = \{w_1, \dots, w_n\}$ ko'rinishida berilgan bo'lsin. Nazorat obyektlari $w_i, i = 1, \dots, n$ ni $X_p, p = \overline{1, r}$ sinflardan qaysi biriga tegishli ekanligini topish talab etiladi. Bunda $x_{p1}, x_{p2}, \dots, x_{pm_p} \in X_p, p = \overline{1, r}$ sinflari obyektlarining belgilari o'xshashligi kattaligi oldindan aniqlangan δ sonidan kam bo'lmasligi talab etiladi. Bu yerda X_p sinf m_p ta x_{p1}, \dots, x_{pm_p} obyektlardan tashkil topgan $X = \bigcup_{p=1}^r X_p$ bo'ladi. Nominal belgilar fazosida obyektlarning belgilari o'xshashligi ko'rsatuvchi kattalik $\rho^j(w_i, x_{pq})$ orqali belgilanib, (1) formula asosida hisoblansin, ya'ni

$$\rho_{pi}^j(w_i, x_{pq}) = \begin{cases} 1, & \text{agar } (w_i^j - x_{pq}^j) = 0; \\ 0, & \text{aks holda} \end{cases} \quad (15)$$

bu yerda $p = \overline{1, r}; i = \overline{1, n}; q = \overline{1, m_p}; j = \overline{1, N}$; Ifodalangan kattalik vektorning

parametrlari bo‘lib, u quyidagi

$$\rho_{pi}(w_i, x_{pq}) = (\rho_{pi}^1(w_i, x_{pq}), \rho_{pi}^2(w_i, x_{pq}), \dots, \rho_{pi}^N(w_i, x_{pq}))$$

ko‘rinishda ifodalanadi. Demak, p sinfdagi ixtiyoriy w_i va x_{pq} obyektlar uchun Bul vektorlar fazosida $\rho_{pi}(w_i, x_{pq}) = (\rho_{pi}^1(w_i, x_{pq}), \rho_{pi}^2(w_i, x_{pq}), \dots, \rho_{pi}^N(w_i, x_{pq}))$ Bul vektoriga ega bo‘linadi. Bu vektorning komponentalari qaralayotgan belgi kesimida ikkita obyektlarining muhim belgilari o‘xshashligini bildiradi. Agar $\rho_{pi}^j(w_i, x_{pq}) = 1$ bo‘lsa, u holda w_i va x_{pq} obyektlar j belgi hisoblash ishlarida ishtirok etadi, aks holda ya’ni $\rho_{pi}^j(w_i, x_{pq}) = 0$ bo‘lsa, u holda shu j belgi hisoblash ishlarida ishtirok etmasligini bildiradi.

Faraz qilaylik, $\Gamma_p(w_i, X_p) = \frac{1}{m_p} \sum_{k=1}^{m_p} (\rho_{pi}(w_i, x_{pk}))$ kattalik w_i nazorat obyektning X_p sinfnining obyektlari bilan o‘rtacha o‘xshashlik darajalari hisoblanishi, xuddi shuningdek, $\Gamma_q(w_i, X_q) = \frac{1}{m_q} \sum_{k=1}^{m_q} (\rho_{qi}(w_i, x_{qk}))$ kattalik w_i obyektning X_q sinfnining barcha obyektlari bilan o‘rtacha o‘xshashlik darajalari hisoblanishi ifodalaydi. Hal qiluvchi qoida sifatida esa aniqlangan baholarni hisoblash algoritmidan foydalanamiz.

$$\mathfrak{R}(W)|_\lambda = \begin{cases} p, \text{ agar } \Gamma_p|_\lambda - \Gamma_q|_\lambda \geq \delta, q \neq p, q = \overline{1, r} \\ 0, \text{ boshqa barcha hollarda} \end{cases} \quad (16)$$

bu yerda $\Gamma_q|_\lambda = \frac{1}{m_p} \sum_{k=1}^{m_p} \rho_{pi}(w_i, x_{pk})|_\lambda$ yaqinlik funksiyasi, δ – oldindan berilgan biror bir musbat son (ishonchlilik ko‘rsatkichi).

Takomillashtirilgan tashxislash algoritmi

Algoritm quyidagi qadamlardan iborat:

1-qadam. Boshlang‘ich ma’lumotlar bazasi barcha

$X_p, p = \overline{1, r}$ sinf obyektlari kesimida shakllantiriladi va nazorat tanlanmasi $W = \{w_1, \dots, w_n\}$ ko‘rinishida berilgan bo‘lsin.

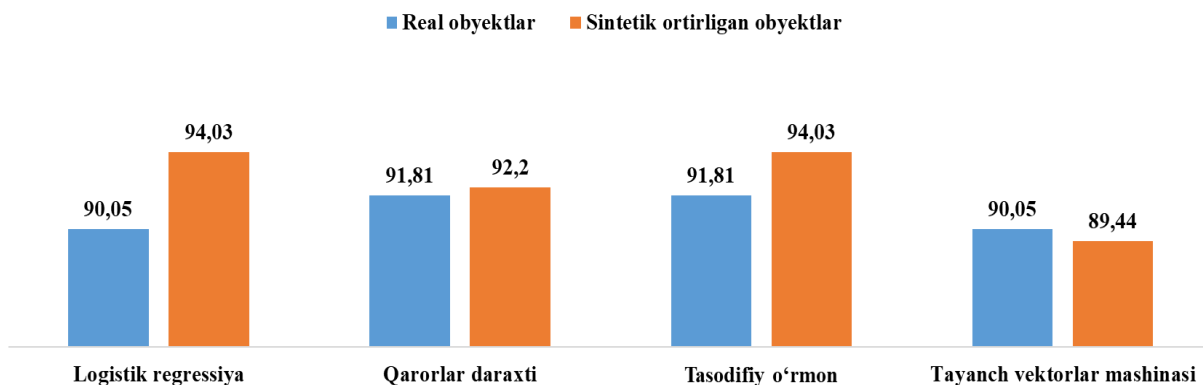
2-qadam. Nazorat obyektlari w_i yuqorida keltirilgan formulalar asosida barcha $p = \overline{1, r}; i = \overline{1, n}; k = \overline{1, m_p}; j = \overline{1, N}$; lar uchun $\rho_{pi}(w_i, x_{pk})$ vektorning barcha parametrlari hisoblanadi.

3-qadam. Shakllantirilgan har bir sinf uchun obyektarning o‘rtacha o‘xshashligini ko‘rsatuvchi kattaligining qiymatlari aniqlanadi.

$$\Gamma_p(w_i, X_p) = \frac{1}{m_p} \sum_{k=1}^{m_p} (\rho_{pi}(w_i, x_{pk}))$$

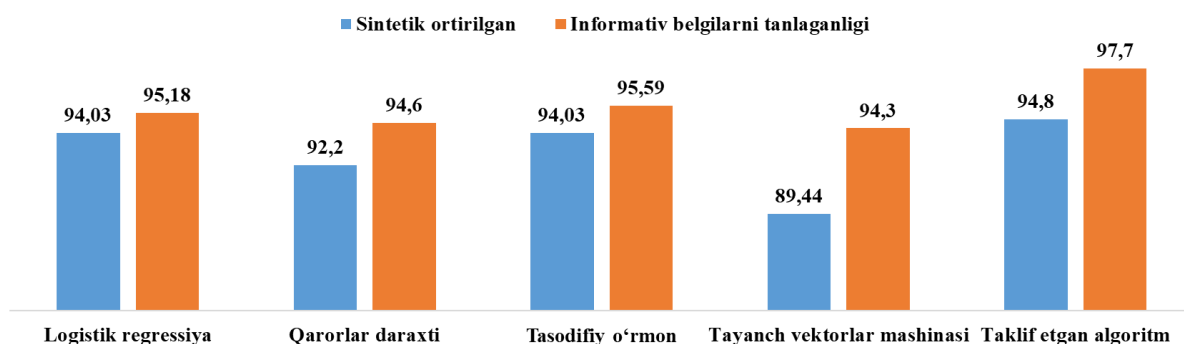
4-qadam. Barcha nazorat obyektlari w_i sinflarga nisbatan o‘xshashligi aniqlanadi. Agar $\Gamma_p|_\lambda - \Gamma_q|_\lambda \geq \delta, q \neq p, q = \overline{1, r}$; sharti bajarilsa p - sinfga tegishli deb qaraladi.

Sintetik obyektlar qo‘shilishi natijasida modellarning aniqlik ko‘rsatkichlari oshgan. Bu esa ma’lumotlar hajmini kengaytirish klasterlash va tasniflash sifatiga ijobiy ta’sir ko‘rsatishini bildiradi(4-rasmda).



4-rasm. Real obyektlar va sintetik obyektlar tahlili

Sintetik obyektlar asosida olingan natijalar hamda informativ belgilarni tanlashdan keyingi ko'rsatkichlar o'zaro taqqoslangan. Olingan natijalar shuni ko'rsatadiki, informativ belgilarni tanlash algoritmi qo'llanilganda barcha modellarda aniqlik darajasi sezilarli oshadi, bu esa taklif etilgan yondashuv samaradorligini tasdiqlaydi(5-rasm).



5-rasm. Sintetik obyektlar tahlili va informative belgilarni tanlaganligi tahlili

Dissertatsiyaning “SBSK” dasturiy majmuasining tibbiyot amaliy masalalarni yechishga tatbiqi” deb nomlangan to'rtinchi bobida dissertatsiya ishi doirasida ishlab chiqilgan algoritmlar asosida yaratilgan “SBSK” dasturiy majmuasining umumiy tuzilmasi, funksional imkoniyatlari va grafik interfeysi hamda ishlash mexanizimi haqida batafsil bayon etilgan.

4.1-paragrafda Veb-platforma ko'rinishida ishlab chiqilgan “Smart Diagnosis” tizimi foydalanuvchilarning tibbiy belgilari va so'rovnomalar ma'lumotlarini tahlil qilgan holda diagnostika olib boradi. Diagnostika qilish orqali bir nechta algoritmlardan iborat mashinali o'qitish, optimal belgi tanlash va tasniflash xatoliklarini minimallashtirish texnikalariga asoslangan. Platforma foydalanuvchisi kiritgan ma'lumotlarni qayta ishlaydi, tashxis natijalarini chiqaradi va tibbiy qarorlarni qo'llab-quvvatlaydi. Tizimning amaliy jihatlari, modellashtirilgan algoritmlari, diagnostika aniqligi va veb-interfeys bilan integratsiyasi batafsil tahlil qilingan. Tadqiqot natijalari ko'rsatadiki, taklif etilgan tizim sut bezi saratonini erta bosqichda yuqori aniqlik bilan aniqlash imkonini beradi va u zamonaviy tibbiy tashxis sohasida katta amaliy ahamiyatga ega.

4.2-paragrafda SBSKning dasturiy majmuasi tibbiy diagnostika sohasida foydalanuvchidan olingan subyektiv simptomatik ma'lumotlarni sun'iy intellekt yordamida tahlil qilish va natijalarni foydalanuvchiga grafik interfeys orqali taqdim

etish imkonini beruvchi innovatsion axborot tizimi yaratildi. Ushbu dasturiy vositaning funksional imkoniyatlari, arxitektura komponentlari, backend tizimi, ma'lumotlar bazasi bilan o'zaro bog'liqligi hamda Python va PHP texnologiyalarining birgalikdagi ishlash mexanizmi yoritiladi.

XULOSA

“Sut bezi saratoni kasalliklarini erta aniqlashda xatoliklarni minimallashtirishga asoslangan tashxislash algoritmi” mavzusidagi dissertatsiya ishi bo'yicha olib borilgan tadqiqotlar natijasida quyidagi xulosalar taqdim etildi:

1. Sut bezi saratoni bemorlarining kasallik varaqalari asosida belgilar majmuasi va umumiy tanlamalarni shakllantirish, ma'lumotlarga dastlabki ishlov berish, obyekt va belgilar muhimligi mezonlari va algoritmlari tahlil qilindi. Sut bezi saratoni klinik ma'lumotlariga dastlabki ishlov berishning adaptiv tozalash algoritmi takomillashtirildi. Mazkur algoritm o'quv tanlamasi va dataset tarkibida mavjud bo'lgan shovqin, kamchilik va noaniqliklarni avtomatik ravishda bartaraf etish imkonini beradi. Natijada datasetdagi ma'lumotlar sifati sezilarli darajada oshirildi hamda keyingi klasterlash va tashxislash bosqichlari uchun ishonchli va tozalangan tanlama asosi yaratildi, bu esa umumiy tashxislash samaradorligini oshirishga muhim hissa qo'shdi.

2. Obyektlarni tashxislashda minimal qamrovli daraxt (MST) yondashuviga asoslangan mikroklasterlash algoritmi ishlab chiqildi. Ushbu algoritm sut bezi saratoni ma'lumotlar to'plamidagi yashirin strukturaviy bog'liqliklarni yuqori o'lchamlilik sharoitida ham aniqlik va samaradorlik bilan aniqlash imkonini beradi. Natijada ishlab chiqilgan dastur shifokorlar uchun qaror qabul qilishni qo'llab-quvvatlovchi vosita sifatida xizmat qilib bemorlarning holatini baholash hamda ularni onkolog va radiolog mutaxassislariga yo'naltirish zaruriyatini aniqlashda yordam berdi. Bu esa soha mutaxassislarining ish unumdorligini o'rtacha 20% ga oshirish imkonini berdi.

3. Obyektlarni sinflashtirish xatolik koeffitsienti va ehtimollar nazariyasiga asoslangan informativ belgilar majmuasini tanlash algoritmi ishlab chiqildi. Sut bezi saratoni kasalliklarini xarakterlovchi 32 ta belgilar majmuasidan tashxis qo'yishda eng muhim bo'lgan 18 ta belgidan iborat informativ belgilar majmuasi tanlab olindi. Mazkur yondashuv tashxislash modelining murakkabligini kamaytirgan holda uning aniqligi va barqarorligini saqlab qolish imkonini berdi hamda kasallikka tashxis qo'yishga ketadigan vaqtni taxminan 20-25% ga qisqartirish imkonini yaratdi.

4. Informativ belgilar fazosida sinflar obyektlari o'xshashlik darajalari farqi va tashxislashdagi xatoliklarni minimallashtirishga asoslangan tashxislash algoritmi ishlab chiqildi. Radiolog va onkologlar tomonidan sut bezi saratoni kasalliklariga tashxis qo'yishda muayyan tibbiy choralarni o'z vaqtida amalga oshirish imkoniyati yaratildi. Klinik sinovlar natijalariga ko'ra tashxis ma'lumotlarni tahlil qilish uchun ketadigan vaqt 20% ga qisqardi va tashxislash samaradorligi 1,7 barobarga oshdi.

5. Tadqiqot natijalari asosida “SBSK” va “SmartDiagnosis” dasturiy majmualari ishlab chiqilib, amaliyotga joriy etildi. Ushbu yechimlar bir qator tibbiy muassasalarda muvaffaqiyatli qo'llanilib, murakkab klinik ma'lumotlarni tahlil qilish uchun universal metodologik asosni shakllantirdi.

**НАУЧНЫЙ СОВЕТ DSc.09/2025.27.12.T.01.03 ПО ПРИСУЖДЕНИЮ
УЧЕНЫХ СТЕПЕНЕЙ ПРИ ТАШКЕНТСКОМ УНИВЕРСИТЕТЕ
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ**

**ДЕНОВСКИЙ ИНСТИТУТ ПРЕДПРИНИМАТЕЛЬСТВА И
ПЕДАГОГИКИ**

ХАЙДАРОВ ШЕРАЛИ ИСЛАМ УГЛИ

**АЛГОРИТМ ДИАГНОСТИКИ, НАПРАВЛЕННЫЙ НА
МИНИМИЗАЦИЮ ОШИБОК ПРИ РАННЕМ ВЫЯВЛЕНИИ РАКА
МОЛОЧНОЙ ЖЕЛЕЗЫ**

05.01.03 - Теоретические основы информатики

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

Ташкент - 2026

Тема диссертации доктора философии (PhD) по техническим наукам зарегистрирована в Высшей аттестационной комиссии при Министерстве высшего образования, науки и инноваций Республики Узбекистан под номером B2025.2.PhD/T5627.

Диссертация выполнена в Деновском институте предпринимательства и педагогики.

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.ziyonet.uz).

Научный руководитель: Мамажанов Рахматилла Якубжанович
кандидат технических наук, доцент.

Официальные оппоненты: Худайбердиев Мирзаакбар Хаккулмирзаевич
доктор технических наук, ст. науч. сотр.
Бобоев Лочинбек Боймуротович
доктор философии (PhD) по техническим наукам

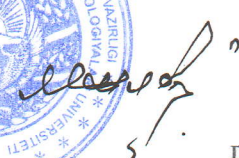
Ведущая организация: Ферганский государственный технический университет

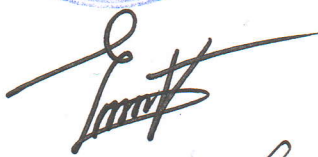
Защита диссертации состоится “10” июня 2026 года в 16⁰⁰ часов на заседании Научного совета DSc.09/2025.27.12.T.01.03 при Ташкентском университете информационных технологий. (Адрес: 100084, г. Ташкент, ул. Амира Темура, д. 108. Тел.: (99871) 207-59-42, e-mail: ilmiy_kengash@tuit.uz).

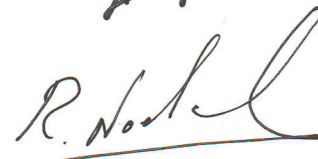
С диссертацией можно ознакомиться в Информационно-ресурсном центре Ташкентском университете информационных технологий имени Мухаммада аль-Хорезми (регистрационный номер № 416). (Адрес: 100084, г. Ташкент, ул. Амира Темура, д. 108. Тел.: (99871) 238-64-15).

Автореферат диссертации разослан “22” мая 2026 года.
(Протокол реестра рассылки № 23 от “22” мая 2026 года.)




М.М.Камилов
Председатель научного совета по присуждению ученых степеней, доктор технических наук, академик АН РУз


Н.А.Эгамбердиев
Ученый секретарь научного совета по присуждению ученых степеней, доктор философии по техническим наукам


Н.О.Рахимов
Председатель научного семинара при научном совете по присуждению ученых степеней, доктор технических наук, доцент

ВВЕДЕНИЕ(аннотация диссертации доктора философии (PhD))

Актуальность и необходимость темы диссертационного исследования. Проведение интеллектуального анализа формирующихся на основе различных источников данных, имеющих большой объём и сложную структуру, выделение из них важных и информативных признаков, а также повышение точности диагностики занимают одно из ведущих мест в развитии теории и практики интеллектуального анализа данных. В наборах данных, применяемых в различных сферах по всему миру, наличие избыточных, повторяющихся и зашумлённых признаков отрицательно сказывается на достоверности результатов классификации, приводя к возникновению ошибок в процессе диагностики. Данное обстоятельство обуславливает необходимость разработки и внедрения в практику эффективных алгоритмических подходов, направленных на предварительную обработку данных, оценку значимости признаков и минимизацию ошибочных решений. В связи с этим приоритетное значение приобретают научные исследования, направленные на сокращение размерности пространства признаков, исключение избыточных атрибутов и формирование совокупности наиболее информативных признаков. Наряду с этим, важную роль играет выявление взаимосвязей между параметрами, характеризующими свойства объектов, оценка степени их влияния, а также использование теоретических подходов и алгоритмического обеспечения, способствующих повышению точности классификации.

В мировом масштабе ведутся научно-исследовательские работы, направленные на оценку взаимной близости объектов и степени их сходства внутри классов на основе многомерных данных, формирующихся из различных источников и имеющих сложную структуру, а также на выделение важных и информативных признаков. В данном направлении, в связи с тем что наличие избыточных, зашумлённых и повторяющихся признаков негативно сказывается на точности и достоверности результатов, возрастает потребность в разработке новых интеллектуальных методов отбора признаков. В частности, особое внимание уделяется созданию алгоритмов, направленных на отбор значимых признаков с помощью методов машинного обучения при определении степени сходства между объектами на основе минимального остоного дерева, а также на снижение показателей ошибок классификации.

В нашей республике последовательно осуществляются реформы по созданию и внедрению в практику алгоритмов, направленных на предварительную обработку данных, интеллектуальный анализ, автоматическую группировку и распознавание объектов, а также аппаратно-программных комплексов, что позволяет достигать определённых научно-практических результатов. В рамках приоритетного направления «Цифровизация экономики и внедрение современных технологий», являющегося одним из ключевых направлений стратегии «Узбекистан-2030», в частности, определена одна из приоритетных задач: «...полная цифровизация отраслей экономики, социальной сферы и системы государственного

управления, увеличение объёмов экспорта ИТ-услуг, а также развитие цифровой инфраструктуры...»¹. В решении этих задач на практике важное значение имеет разработка систем интеллектуального принятия решений на основе больших объёмов данных, формирующихся в сферах медицины, образования, промышленности и сельского хозяйства, а также создание алгоритмов машинного обучения, направленных на выбор значимых признаков, основанных на минимизации ошибок в процессах диагностики, повышение точности классификации и снижение показателей ошибок.

Указ Президента Республики Узбекистан № УП-60 от 28 января 2022 года «О Стратегии развития нового Узбекистана на 2022–2026 годы», Указ № УП -6097 от 29 октября 2020 года «Об утверждении Концепции развития науки до 2030 года», Постановление № ПП-358 от 14 октября 2024 года «Об утверждении Стратегии развития технологий искусственного интеллекта до 2030 года», Постановление Кабинета Министров Республики Узбекистан № 280 от 29 апреля 2025 года «О мерах по дальнейшему развитию Республиканского специализированного научно-практического медицинского центра эндокринологии, а также услуг детской онкологии, гематологии и иммунологии», Постановление № ПП-415 от 28 декабря 2023 года «О дополнительных мерах по ускорению цифровизации системы здравоохранения и внедрению передовых цифровых технологий», а также другие нормативно-правовые документы, относящиеся к данной деятельности, реализация задач, определённых в указанных документах, в определённой степени обеспечивается данным диссертационным исследованием.

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

Степень изученности проблемы. В последние годы научно-практические исследования в области раннего выявления и диагностики рака молочной железы активно развиваются, разрабатываются различные подходы, направленные на обнаружение заболевания на начальных стадиях и повышение точности диагностики. В данном направлении особое внимание уделяется формированию комплекса признаков заболевания, созданию обучающих выборок, оценке состояния пациентов на основе медицинских данных, а также разработке и совершенствованию моделей, методов и алгоритмов, предназначенных для диагностики.

В области раннего выявления рака молочной железы, разработки диагностических моделей и анализа на основе клинических данных проводили исследования зарубежные учёные: Р.Барзилай, К.Леман, К.Ланг, С. Маккинни, Г.Сирас, М. Френк, Ли Сюань, Ш. Хусейн, Дж. Смит, А. Лопес, К. Вонг и

¹ Указ Президента Республики Узбекистан № УП -158

другие. В нашей республике весомый вклад в развитие теоретических основ анализа данных внесли учёные М.М. Камилов, Ш.Х. Фозилов, Ф.Т.Адилова, А.Х. Нишанов, Н.С. Маматов, Б.Б. Акбараалиев и другие. Кроме того, в направлении разработки алгоритмов и программных средств, направленных на диагностику и прогнозирование рака молочной железы, исследовательские работы проводили Д.Т. Мухамедиева, О.Б. Рузибаев и другие.

Анализ научных работ в данном направлении показывает, что в существующих исследованиях основное внимание уделяется построению диагностических моделей, при этом алгоритмические подходы, направленные на формирование набора признаков с высокой информативностью, чётко разделяющих классы из исходного пространства признаков, исключение избыточных и малозначимых признаков, а также минимизацию ошибок диагностики с учётом степени сходства между объектами, изучены в недостаточной степени.

Связь исследования с планами научно-исследовательских работ высшего учебного или научно-исследовательского учреждения, в котором выполнена диссертация. В рамках диссертационного исследования № SICA-UZ-ISG-2022-004 «Taskin» в информационном центре по защите социальных, экономических и правовых интересов женщин Денауского района Сурхандарьинской области, в рамках программы социальных инноваций в Центральной Азии, финансируемой Агентством международного развития Соединённых Штатов Америки (USAID), была разработана платформа обращений «Depov404».

Цель исследования. заключается в разработке алгоритма минимизации диагностических ошибок при раннем выявлении рака молочной железы

Задачи исследования:

анализ методов и алгоритмов предварительной обработки клинических данных по заболеваниям рака молочной железы;

разработка алгоритма кластеризации, основанного на подходе минимального остовного дерева, для диагностики объектов;

разработка алгоритма выбора совокупности информативных признаков, основанного на коэффициенте ошибок классификации объектов и теории вероятностей;

разработка алгоритма диагностики, основанного на минимизации разницы степеней сходства объектов классов в пространстве информативных признаков и ошибок диагностики;

разработка и внедрение в практику программы для ранней диагностики заболеваний рака молочной железы в медицине.

Объект исследования. процессы предварительной обработки медицинских данных, кластеризации и минимизации ошибок при диагностике объектов в интеллектуальном анализе данных.

Предмет исследования. предварительная обработка данных, кластеризация на основе микрокластеризации, выбор совокупности информативных признаков, а также алгоритмы и программные средства диагностики объектов, основанные на минимизации ошибок.

Методы исследования. В процессе исследования использовались методы предварительной обработки медицинских данных, микрокластеризации, направленной на повышение точности эталонов, коэффициента ошибок и теории вероятностей при выборе информативных признаков, а также методы ранней диагностики.

Научная новизна исследования заключается в следующем:

усовершенствован на основе теории вероятностей алгоритм адаптивной очистки для предварительной обработки, направленный на выявление и устранение зашумленных, повторяющихся, отсутствующих и несогласованных значений, встречающихся в клинических данных по раку молочной железы;

разработан алгоритм микрокластеризации на основе минимального остовного дерева, направленный на повышение точности группировки с учётом степеней близости и сходства при диагностике объектов;

разработан алгоритм выбора совокупности информативных признаков, направленный на оценку степени значимости признаков на основе коэффициента ошибок и теории вероятностей, а также выделение наиболее важных характеристик при классификации объектов;

разработан алгоритм диагностики, направленный на определение разницы степеней сходства объектов классов в пространстве информативных признаков, минимизацию ошибок, возникающих в процессе диагностики, а также повышение точности принятия решений.

Практические результаты исследования заключаются в следующем:

разработана программа в рамках многоэтапного усовершенствованного алгоритма адаптивной очистки и устранения шумов для предварительной обработки медицинских данных;

разработан программный инструмент на основе алгоритмов микрокластеризации и классификации, основанных на минимальном остовном дереве, для кластеризации объектов;

разработана программа, включающая алгоритм выбора информативных признаков, направленный на минимизацию ошибок при ранней диагностике рака молочной железы;

Достоверность результатов исследования. Достоверность предлагаемых алгоритмов минимизации ошибок подтверждается разработкой методов адаптивной очистки и алгоритмов устранения зашумленных признаков для повышения качества данных, статистических критериев и коэффициентов ошибок для обеспечения эффективности выбора информативных признаков, а также алгоритмов микрокластеризации на основе минимального остовного дерева и классификации объектов, снижения ошибок и повышения точности результатов. При этом многоэтапный подход гарантирует методологическую обоснованность исследования и достоверность полученных результатов.

Научная и практическая значимость результатов исследования.

Научная значимость результатов исследования объясняется совершенствованием адаптивного алгоритма очистки при предварительной

обработке клинических данных рака молочной железы, разработкой алгоритма микрокластеризации на основе минимального остовного дерева, а также созданием методов отбора набора информативных признаков на основе коэффициента ошибок и теории вероятностей. Кроме того, разработка алгоритма диагностики, основанного на различии степеней сходства между объектами в пространстве информативных признаков, позволяет повысить точность классификации и снизить сложность модели, что повышает эффективность процесса анализа данных и характеризуется научной значимостью полученных результатов.

Практическая значимость результатов исследования обусловлена созданием программного комплекса «SmartDiagnosis», предназначенного для ранней диагностики рака молочной железы на основе разработанных алгоритмов. Данное программное средство служит для анализа клинических и маммографических данных в медицинских учреждениях, повышения точности диагностики и поддержки врачей в процессе принятия решений, что позволяет снизить диагностические ошибки и повысить надёжность результатов, что подтверждает его практическую значимость.

Внедрение результатов исследования. На основе результатов, полученных при использовании разработанного в исследовании алгоритма диагностики, основанного на минимизации ошибок при раннем выявлении заболеваний рака молочной железы, и программного комплекса «Smartdiagnosis»:

разработан и внедрён в практику с целью совершенствования процессов диагностики в Республиканском специализированном онкологическом и радиологическом научно-практическом медицинском центре, Сурхандарьинский филиал (г. Термез, Сурхандарьинская область, Республика Узбекистан) алгоритм адаптивной очистки для предварительной обработки клинических данных по раку молочной железы, основанный на теории вероятностей, реализованный в программе «SBSK» (справка Министерства здравоохранения Республики Узбекистан от 5 ноября 2025 года № 02-28/26482). В результате из 32 симптоматических признаков, описывающих рак молочной железы, отобрано 18 наиболее значимых информативных признаков для диагностики, благодаря чему время, затрачиваемое на постановку диагноза на основе данных признаков, сокращено на 20–25%, а достоверность диагностики повышена на 20%;

разработан и внедрён в практику с целью совершенствования процессов диагностики и поддержки принятия решений в Термезском филиале Ташкентского государственного медицинского университета (г. Термез, Сурхандарьинская область, Республика Узбекистан) алгоритм микрокластеризации на основе минимального остовного дерева для диагностики объектов по степеням близости и сходства, реализованный в программе «Smartdiagnosis» (справка Министерства здравоохранения Республики Узбекистан от 5 ноября 2025 года № 02-28/26482). В результате с использованием отобранного набора информативных признаков время,

затрачиваемое на диагностику заболевания, сокращено в 1,5 раза, а эффективность диагностики повышена в 1,7 раза;

разработан и внедрён в практику с целью повышения эффективности диагностики в диагностическо-лечебном центре «Утан Полвон» (г. Термез, Сурхандарьинская область, Республика Узбекистан) алгоритм диагностики, направленный на снижение коэффициента ошибок в пространстве информативных признаков при минимизации ошибок в диагностике объектов, реализованный в программе «Smartdiagnosis» (справка Министерства здравоохранения Республики Узбекистан от 5 ноября 2025 года № 02-28/26482). В результате на основе разработанных алгоритмов автоматизированы процессы диагностики рака молочной железы, а также подтверждено повышение достоверности диагностического процесса на 20%.

Апробация результатов исследования. Результаты исследований обсуждались на 5 международных и 3 республиканской научно-практической конференциях.

Опубликованность результатов исследования. По теме диссертации в общей сложности по теме исследования опубликовано 20 научных работ, из которых 9 статей опубликованы в научных изданиях, рекомендованных к публикации Высшей аттестационной комиссией Республики Узбекистан, 6 в зарубежных и 3 в республиканских журналах, также получено 3 свидетельство об официальной регистрации программы для ЭВМ.

Структура и объём диссертации. Диссертация состоит из введения, четырёх глав, заключения, списка использованной литературы и приложений. Объём диссертации составляет 118 страниц.

ОСНОВНОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ

Во введении обоснованы актуальность и необходимость темы диссертации, показано её соответствие приоритетным направлениям развития науки и технологий республики, степень изученности темы, а также связь диссертации с планами научно-исследовательских работ высшего образовательного учреждения, в котором она выполнена; изложены цель, задачи, объект, предмет и методы исследования; раскрыты научная новизна и практические результаты исследования, достоверность результатов, научная и практическая значимость полученных результатов; приведены сведения о внедрении результатов исследования, их апробации, опубликованных работах и структуре диссертации.

В первой главе диссертации, озаглавленной «**Анализ методов и алгоритмов раннего выявления рака молочной железы**», изложено, что методы машинного обучения в современных технологиях обеспечивают достаточно высокую точность и эффективность при выявлении онкологических заболеваний. Данные технологии позволяют анализировать медицинские данные, определять диагностические признаки заболевания и автоматизировать процессы принятия решений.

В 1.1 приведены подробные аналитические сведения об этапах подготовки симптомов рака молочной железы для машинного обучения,

научных основах понятия информационной значимости признаков, а также об информативной значимости клинических признаков.

В 1.2 проанализированы современные подходы к раннему выявлению онкологических заболеваний, значение раннего выявления рака в повышении эффективности лечения, а также возможности постановки высокоточного диагноза на основе визуальной диагностики, биомаркеров и генетических данных с использованием современных алгоритмов искусственного интеллекта.

В 1.3 изучены и обоснованы критерии и алгоритмы выявления онкологических заболеваний, клинические, лабораторные, визуальные, гистологические и биомаркерные критерии, применяемые при выявлении рака, а также современные диагностические подходы, основанные на алгоритмах искусственного интеллекта и машинного обучения.

В 1.4 изложена постановка основной задачи диссертационной работы, приведена постановка оптимизационной задачи, основанной на минимизации диагностических ошибок, а также подходы к её решению.

Предположим, что обучающая выборка X задана в следующем виде, то есть

$$X = \bigcup_{p=1}^r X_p, X_p \cap X_q = \emptyset, p \neq q, p, q = \overline{1, r}, \\ X_p = \{x_{pi} = (x_{pi}^1, x_{pi}^2, \dots, x_{pi}^N): i = \overline{1, m_p}\} \quad (1)$$

где r - число классов, X_p - классы ($p = \overline{1, r}$), m_p - число объектов в классе X_p . При формировании оптимальных информативных систем, описывающих объекты, используются некоторые понятия, связанные с вектором информативности $\lambda = (\lambda^1, \lambda^2, \dots, \lambda^N)$. Относительно заданного вектора λ информативную систему, описывающую объекты, определим следующим образом(2).

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

Определение. Если выполняется условие $\sum_{j=1}^N \lambda^j = \ell$ то вектор $\lambda = (\lambda^1, \lambda^2, \dots, \lambda^N)$ называется ℓ - информативным. Множество всех ℓ информативных векторов обозначим через Λ^{ℓ} .

$$\Lambda^{\ell} = \{\lambda: \sum_{j=1}^N \lambda^j = \ell, \lambda^j \in \{0,1\}, j = \overline{1, N}\} \quad (3)$$

Как видно из выражения (3), мощность множества Λ^{ℓ} равна $|\Lambda^{\ell}| = C_N^{\ell} = \frac{N!}{\ell!(N-\ell)!}$ тогда $\Lambda^{\ell}: X \rightarrow X|_{\Lambda^{\ell}} = \{X|_{\lambda} | \lambda \in \Lambda^{\ell}\}$.

Кроме того, пусть задано множество l -информативных векторов Λ^l . тогда информативные системы, описывающие объекты, соответствующие произвольному вектору $\lambda \in \Lambda^l$ определяются следующим образом:

$$X|_{\lambda} = \{x \in X: x = (\lambda^1 x^1, \lambda^2 x^2, \dots, \lambda^N x^N), \sum_{i=1}^N \lambda^i = \ell\}. \quad (4)$$

В интеллектуальном анализе данных важным является определение поведения объектов обучающей выборки в пространстве признаков различной размерности, то есть при переходе из пространства признаков одной размерности в пространство признаков другой размерности. Например, при решении задачи классификации в результате изменения размерности

пространства признаков некоторые объекты могут быть отнесены не к своим классам, а перейти в другой класс, то есть

$\exists x \in X, \lambda \in \Lambda^{\ell_1}, \mu \in \Lambda^{\ell_2}, \ell_1 \neq \ell_2$ такие, что $x|_{\lambda} \in X_p, x|_{\mu} \in X_q, p \neq q$.

С целью выявления таких ситуаций и их предотвращения при уменьшении размерности пространства признаков учитываются ошибки классификации. Предположим, что для l -информативного вектора относительно вектора λ коэффициент ошибок классификации равен $\theta(l)|_{\lambda}$, а число объектов, не распознавших свой класс, то есть неправильно классифицированных, равно $\kappa(l)|_{\lambda}$. Связь между коэффициентом ошибок классификации и числом неправильно классифицированных объектов определяется следующим образом.

$$\theta(\ell)|_{\lambda} = \frac{\kappa(\ell)|_{\lambda}}{m} \quad (5)$$

здесь $m = \sum_{p=1}^r m_p$ общее количество объектов в обучающей выборке.

Примечание. Если важно, чтобы объекты определённого класса (классов) не были ошибочно отнесены к другому классу, а также наоборот чтобы объекты других классов не попадали ошибочно в данный класс (классы), то при определении коэффициента ошибок классификации необходимо учитывать степень важности классов. Например, пусть степень значимости недопущения ошибочного отнесения объектов класса X_p к другим классам равна α_p , а степень значимости недопущения ошибочного отнесения объектов других классов к данному классу равна β_p . Тогда количество неправильно классифицированных объектов для класса X_p с учётом степени значимости может быть определено следующим образом.

$$\kappa(\ell)|_{\lambda}(X_p) = \alpha_p \cdot t_p^{out}|_{\lambda} + \beta_p \cdot t_p^{in}|_{\lambda} \quad (6)$$

здесь t_p^{out} и t_p^{in} - соответственно количество объектов, ошибочно выбывших из класса X_p и ошибочно включённых в класс X_p . На основании выражения (5) формула (6) принимает следующий вид:

$$\theta(\ell)|_{\lambda} = \frac{\sum_{p=1}^r (\alpha_p \cdot t_p^{out}|_{\lambda} + \beta_p \cdot t_p^{in}|_{\lambda})}{m} \quad (7)$$

Основной целью диссертационного исследования является построение такого решающего правила, при котором ошибка классификации была бы минимальной, то есть

$$\begin{cases} \theta(\ell)|_{\lambda} = \frac{\kappa(\ell)|_{\lambda}}{m} \rightarrow \min \\ \Lambda^{\ell} = \{ \lambda: \sum_{j=1}^N \lambda^j = \ell, \lambda^j \in \{0,1\}, j = \overline{1, N} \} \end{cases} \quad (8)$$

Во второй главе диссертации, озаглавленной «**Алгоритмы, основанные на минимизации ошибок при классификации медицинских объектов**», на основе алгоритмов адаптивной очистки медицинских данных по раку молочной железы сформирована общая выборка. Также обоснованы возможности реализации процессов микро- и макрокластеризации на основе усовершенствованного алгоритма кластеризации MSTL.

В 2.1 на основе алгоритма адаптивной очистки медицинских данных по раку молочной железы из общей выборки были очищены ошибочно внесённые

признаки внутри объектов, и на основе вероятностного подхода путём выбора наиболее часто встречающегося значения признака сформирована новая общая выборка. Также приведены «Признаки рака молочной железы и принимаемые ими значения» (таблица 1).

Таблица 1.

Признаки рака молочной железы и их принимаемые значения.

Признаки	Симптомы онкологии молочной железы (наименование признаков)	Возможные значения признаков
x^1	В каком возрасте у Вас началось заболевание?	1. Заболевание отсутствует 2. В возрасте 17–30 лет 3. В возрасте 30–50 лет 4. После 50 лет
x^2	Есть ли у Вас образование (опухоль) в молочной железе?	Нет Да
x^3	Наблюдается ли у Вас общая слабость (астения)?	1. Нет 2. Лёгкая (умеренная) 3. Выраженная
...		
x^{32}	Принимаете ли Вы регулярно с целью предохранения контрацептивные или терапевтические гормональные препараты?	Нет Да

В рамках алгоритма адаптивной очистки медицинских данных по раку молочной железы клинические признаки оценивались на основе порядковых номинальных значений, при этом данные признаки были предварительно закодированы врачами-экспертами с опорой на клинический опыт и практику. Сформированная в исследовании общая обучающая выборка основана на реальных клинических наблюдениях, где каждая строка представляет отдельного пациента, а каждый столбец определённый клинический признак, и служит основной базой данных для последующего интеллектуального анализа и алгоритмов кластеризации.

Алгоритм адаптивной очистки реализуется в следующих шагах.

шаг 1. Ввод исходных данных

$$X = \{x_1, x_2, \dots, x_m\}, x_i = (x_i^1, \dots, x_i^N)$$

шаг 2. Контроль допустимых значений. Каждое значение номинального признака проверяется на соответствие клинической логике

$$\tilde{x}_i^j = \begin{cases} x_i^j, & \text{если } x_i^j \in \Omega_j, \\ \emptyset, & \text{в противном случае} \end{cases} \quad (9)$$

шаг 3. Адаптивная очистка и заполнение пропущенных значений Для каждого признака j определяется наиболее часто встречающееся значение:

$$\hat{v}_j = \arg \max_{v \in \Omega_j} P(x^j = v) \quad (10)$$

где

$$P(x^j = v) = \frac{\text{count}(x^j=v)}{m} \quad (11)$$

Пропущенные значения не оставляются пустыми и заполняются по следующему правилу

$$x_i^j = \begin{cases} \tilde{x}_i^j, & \tilde{x}_i^j \neq \emptyset, \\ \hat{v}_j, & \text{в противном случае} \end{cases} \quad (12)$$

что также представляет собой контрольный механизм, основанный на клинической логике.

шаг 4. Формирование итоговой выборки. В результате этапов адаптивной очистки и оценки степени сходства была сформирована общая выборка (рисунок 1).

$$X = \tilde{X} = \{\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_m\} \quad (13)$$

№	В каком возрасте у вас началось заболевание?	Есть ли у вас образование (опухоль) в молочной железе?	Наблюдается ли у вас общая слабость (утомляемость)?	Как давно, по вашему мнению, развивается заболевание?	Были ли случаи заболевания у ваших родственников (мать, сёстры, тётки)?	Есть ли у вас боли в нижней части плечевой области?	Есть ли увеличенные лимфатические узлы в подмышечной области?	Какой размер вашей молочной железы?	...	В каком возрасте вы вышли замуж?	В каком возрасте у вас появился первый ребёнок?	Наблюдаются ли изменения в вашем весе?	Ощущаете ли вы узловое образование в молочной железе?	Подвижно ли образование (опухоль) при пальпации?	Чёткие ли границы у образования в молочной железе?	Каково состояние опухоли в молочной железе?	Принимаете ли вы регулярно гормональные контрацептивные препараты в целях предохранения?
	X1	X2	X3	X4	X5	X6	X7	X8	...	X25	X26	X27	X28	X29	X30	X31	X32
1	4	2	2	2	1	2	2	2	...	2	2	3	2	1	2	1	1
2	3	2	2	2	1	1	1	1	...	3	3	1	1	2	2	1	2
3	3	2	2	3	1	1	2	2	...	1	1	1	1	2	2	1	1
:	:	:	:	:	:	:	:	:	...	:	:	:	:	:	:	:	...
:	:	:	:	:	:	:	:	:	...	:	:	:	:	:	:	:	...
563	3	2	3	2	1	1	2	1	...	2	2	1	2	1	2	3	1
564	4	2	3	2	1	1	2	2	...	2	2	1	2	2	2	3	1
565	3	2	3	3	1	3	1	1	...	2	2	1	2	2	1	3	1
566	3	2	3	2	1	1	2	2	...	1	1	1	2	1	1	3	1
567	3	2	3	2	1	1	2	1	...	2	2	3	2	1	1	3	1

Рисунок 1. Общая выборка

Предложенный алгоритм адаптивной очистки направлен на повышение качества номинальных клинических данных по раку молочной железы без необходимости выполнения дополнительных операций кодирования. Данный подход обеспечивает устранение неопределённостей и пропущенных значений в данных, что способствует повышению точности и устойчивости процессов кластеризации и классификации в высокоразмерном номинальном пространстве.

В 2.2 на основе алгоритма микроагрегации, основанного на минимальном остовном дереве (MST), данные разделяются на группы, содержащие не менее

k элементов. Каждая группа представляется репрезентативным значением, определённым на основе степени сходства между элементами внутри группы. Процесс формирования групп основывается на критерии минимизации потерь информации. На основе структуры MST определяются наиболее близкие связи между объектами.

Усовершенствованный алгоритм кластеризации MSTL.

шаг 1. Предположим, что обучающая выборка X задана в следующем виде, то есть

$$X = \bigcup_{p=1}^r X_p, X_p \cap X_q = \emptyset, p \neq q, p, q = \overline{1, r}, \quad (14)$$

$$X_p = \{x_{pi} = (x_{pi}^1, x_{pi}^2, \dots, x_{pi}^N): i = \overline{1, m_p}\},$$

здесь r - число классов, X_p - классы ($p = \overline{1, r}$), m_p - число объектов в классе X_p . x_{pi} - i -й объект p -го класса. x_{pi}^j - j -й признак i -го объекта p -го класса.

Для решения задачи кластеризации введена функция сходства между объектами.

шаг 2. В пространстве номинальных признаков величина, выражающая сходство объектов, задаётся следующим образом

$$\rho_{pi}^j(x_{pi}, x_{pq}) = \begin{cases} 1, & \text{agar } x_{pi}^j = x_{pq}^j \\ 0, & \text{aks holda} \end{cases} \quad (15)$$

здесь $p = \overline{1, r}; i \neq q = \overline{1, m_p}; j = \overline{1, N}$ Данная величина выражает сходство объектов по каждому признаку. Общее сходство между объектами определяется как среднее арифметическое этих значений.

$$\rho_{pi}(x_{pi}, x_{pq}) = (\rho_{pi}^1(x_{pi}, x_{pq}), \rho_{pi}^2(x_{pi}, x_{pq}), \dots, \rho_{pi}^N(x_{pi}, x_{pq}))$$

$$\rho(x_{pi}, x_{pq}) = \frac{1}{N} \sum_{j=1}^N \rho_{pi}^j(x_{pi}, x_{pq}) \quad (16)$$

шаг 3. Для каждого объекта x_{pi} определяются соседи в пределах порога сходства r_p . Значение радиуса r_p определяется на основе среднего значения показателей сходства, соответствующих k ближайшим соседям объектов класса X_p .

$$r_p = \frac{1}{m_p} \sum_{i=1}^{m_p} \rho^{(k)}(x_{pi}), K_i = \{x_{pq} \in X \mid \rho_{pi}(x_{pi}, x_{pq}) \geq r_p\} \quad (17)$$

Если выполняется условие

$$|K_i| \geq \text{MinPts}, \quad (18)$$

то объект x_{pi} является центром микрокластера. $|K_i| \geq k \Rightarrow x_{pi}$ «центр микрокластера» $\rho^{(k)}(x_{pi})$ - значение общего сходства, соответствующее k -му ближайшему соседу i -го объекта p -го класса x_{pi} , m_p - число объектов в классе $|K_i|$ -количество соседей, соответствующих данному центру.

шаг 4. В процессе микроагрегации для оценки эффективности кластеризации в качестве функций потерь используются показатели SSE (sum of squared errors) и SST (total sum of squares). Данные показатели позволяют

количественно оценить уровень разброса объектов внутри кластеров и в целом.

$$\sum_{p=1}^r m_p = M, SSE = \frac{1}{M} \sum_{p=1}^r \sum_{i=1}^{m_p} (\rho_{pi}(x_{pi}, \bar{x}_p))^2, SST = \frac{1}{m_p} \sum_{i=1}^{m_p} \rho_{pi}(x_{pi}, \bar{x})^2 \quad (19)$$

SSE - мера внутреннего разброса объектов внутри кластера относительно центра кластера.

SST - мера общего разброса всех объектов относительно общего центра. x_{pi} - i -й объект p -го класса; \bar{x}_p - центральное (среднее) значение p -го класса; \bar{x} - общее среднее значение всех объектов; m_p - число объектов в p -м классе.

шаг 5. Показатель L является мерой оценки качества кластеризации и выражает отношение SSE и SST.

$$L = \frac{SSE}{SST} \rightarrow \max \quad (20)$$

Уменьшение значения L свидетельствует о повышении точности кластеризации. При $L = 1$, достигается идеальная кластеризация.

На основе разработанного алгоритма реализуется процесс кластеризации. В результате формируется диаграмма распределения количества объектов по классам (рисунок 2).

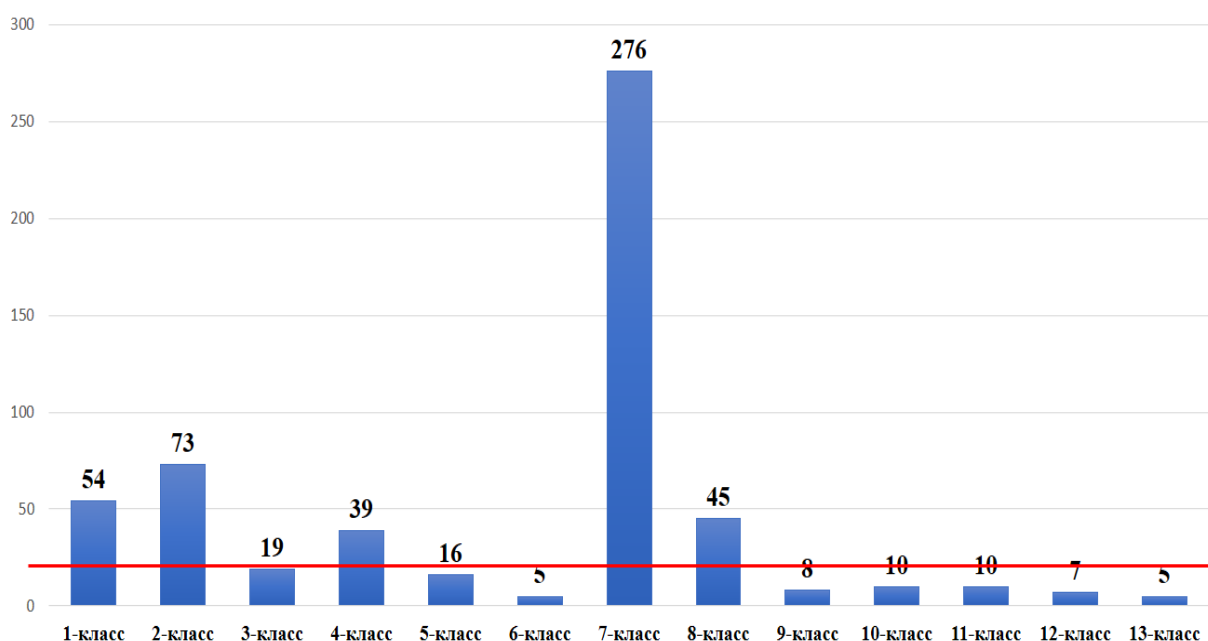


Рисунок 2. Распределение количества объектов по классам

В 2.3 разработан алгоритм формирования синтетических объектов при нормализации обучающей выборки по раку молочной железы, а также критерии, определяющие их содержание, и алгоритм создания полусинтетической обучающей выборки. Данный алгоритм направлен на формирование синтетических объектов в пространстве номинальных признаков путём оценки объектов и создание нового расширенного набора данных.

Алгоритм формирования синтетической обучающей выборки

шаг 1. Объекты общей обучающей выборки вводятся в базу данных. Исходная база данных формируется по всем объектам $x_i \in X, i = \overline{1, M}$

шаг 2. Выполняется процесс предварительной обработки данных. При этом осуществляется очистка данных, поскольку в собранной информации могут присутствовать ошибки, пропущенные значения, дубликаты или неопределённые данные. Указанные недостатки устраняются, после чего данные нормализуются.

шаг 3. Общая обучающая выборка предварительно разделяется на классы $x_{p1}, x_{p2}, \dots, x_{pt_p} \in X_p, p = \overline{1, r}$

шаг 4. Для каждого сформированного класса определяются значения средней сходства объектов $\Gamma_p(X_p), p = \overline{1, r}$

шаг 5. Новый объект \bar{x} для классов $X_p, p = \overline{1, r}$ формируется с сохранением значимых признаков объектов данного класса, обеспечивающих увеличение средней степени сходства $\Gamma_p(X_p)$.

шаг 6. Если в p -м классе количество объектов равно $t_p = 1$, то при формировании синтетического объекта сохраняются признаки данного объекта, отличающие его от объектов других классов.

шаг 7. Шаги 5 и 6 повторяются до тех пор, пока в каждый класс не будет добавлено H новых синтетических объектов. В результате формируются классы, состоящие из гибридных объектов.

В третьей главе диссертации, озаглавленной «Алгоритмы и инновационные подходы раннего выявления рака молочной железы», показано, что отбор набора информативных признаков в системах классификации и прогнозирования имеет важное значение для минимизации ошибок и позволяет повысить точность и обобщающую способность модели за счёт исключения малозначимых и шумовых признаков.

В 3.1 рассматривается задача снижения ошибок классификации путём совершенствования алгоритма отбора набора информативных признаков, основанного на минимизации ошибок. При этом предусматривается решение таких задач, как снижение размерности данных, повышение точности диагностики и прогнозирования, а также улучшение вычислительной эффективности. Кроме того, проводится анализ алгоритмов, применяемых в практической деятельности, и сопоставляется их влияние на уровень ошибок.

Алгоритм отбора информативных признаков, направленный на минимизацию ошибок, реализуется в следующих шагах.

шаг 1. Исходные данные: X - обучающая выборка; X_p ($p = \overline{1, r}$) - классы; r число классов; $k, \ell, N, \theta, T, i, p, \mu$.

шаг 2. Для $\lambda \in \Lambda^N$ вычисляются $\kappa(N)$ и $\theta = \theta(N) |_{\lambda}$

шаг 3. $i = k$

шаг 4. При $v = 1$ для $p_v = (p_v^1, p_v^2, \dots, p_v^N)$ присваиваются значения

$$p_v^j = \frac{1}{N}; j = \overline{1, N}$$

шаг 5. Для всех $\mu_i = \left\{ \underbrace{0,0, \dots, 0}_{N-i}, \underbrace{1,1, \dots, 1}_i \right\}$, $\mu \frac{N!}{i!(N-i)!} = \left\{ \underbrace{1,1, \dots, 1}_i, \underbrace{0,0, \dots, 0}_{N-i} \right\}$

векторов $\mu \in L^i$ вычисляются $\theta(N-i) |_{\lambda-\mu}$ и $\theta(i) |_{\mu}$

шаг 6. Для всех векторов μ , если выполняется условие $\theta(N-i) |_{\lambda-\mu} \leq \theta$, то присваивается $p_{v+1}(\mu) = p_v + \frac{(p_v, \mu)}{N-i} \cdot p_v |_{\lambda-\mu} - p_v |_{\mu}$.

шаг 7. Вычисляется $p_{v+1} = \frac{1}{T} \sum_{\mu} p_{v+1}(\mu)$, где T - число $p_{v+1}(\mu)$. Для вектора $p_v = (p_v^1, p_v^2, \dots, p_v^N)$ проводится проверка и в соответствии с этим i признаков исключаются из пространства признаков.

шаг 8. Присваиваются новые значения $N = N - i$ и $\lambda = (1,1, \dots, 1) \in L^N$

шаг 9. Если $i < N - l$, то осуществляется возврат к шагу 3.

шаг 10. Если $i > N - l$ то принимается $i = N - l$ и осуществляется возврат к шагу 4.

шаг 11. Выходной параметр λ , оптимальный вектор.

При поэтапном увеличении числа признаков с 1 до 18 ошибка классификации существенно снижалась, что свидетельствует об улучшении качества обучения модели по мере получения новой информации. На уровне 18 признаков ошибка достигла минимального значения 2,3% (17 объектов), и данная точка была определена как оптимальная конфигурация. В диапазоне от 19 до 32 признаков ошибка не уменьшалась, а, напротив, постепенно возрастала, что является классическим проявлением эффекта переобучения при включении в модель избыточных и шумовых признаков. Таким образом, из 32 номинальных признаков, представленных на рисунке 3, только 18 были определены как действительно информативные для обеспечения точности модели.

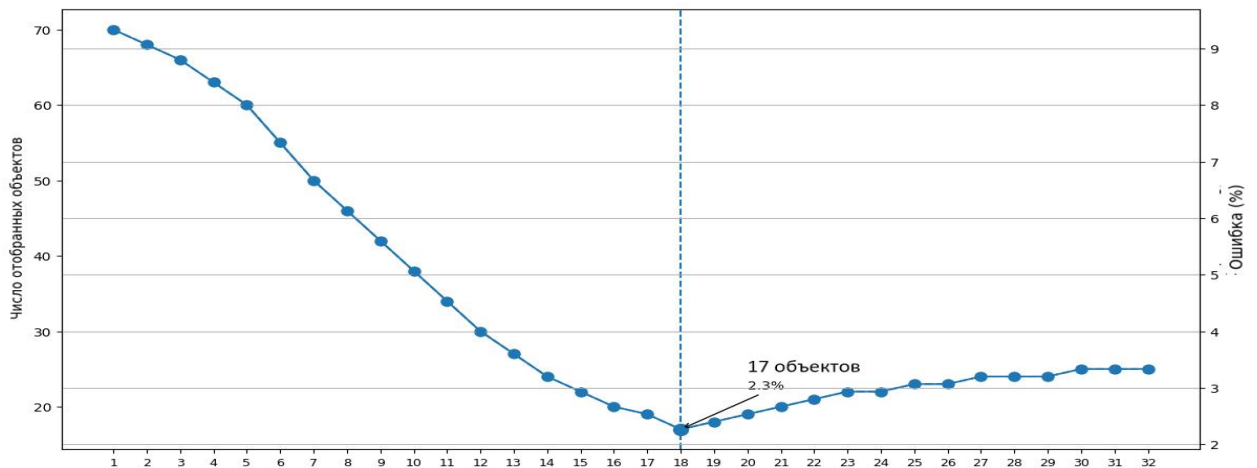


Рисунок 3. Зависимость изменения ошибки классификации от l

В параграфе 3.2 рассматривается диагностический алгоритм, ориентированный на минимизацию ошибок. Пусть в N -мерном пространстве номинальных признаков задана обучающая выборка по заболеванию рака молочной железы $x_{p1}, x_{p2}, \dots, x_{pm_p} \in X_p, p = \overline{1, r}$. Каждый объект обучающей выборки $x_{pi} = (x_{pi}^1, x_{pi}^2, \dots, x_{pi}^N)$, $i = \overline{1, m_p}$ представляет собой пациента, заданного в N -мерном пространстве номинальных признаков. Здесь x_{pi} читается как i -й объект p -го класса, N обозначает число признаков, характеризующих

объекты, а m_p - количество объектов в p -м классе. Таким образом, пациенты $x_{pi} = (x_{pi}^1, x_{pi}^2, \dots, x_{pi}^N)$ образуют $i(i = \overline{1, m_p})$ объекты p -го класса.

Задача. Пусть контрольная выборка задана в виде $W = \{w_1, \dots, w_n\}$. Требуется определить, к какому из классов $X_p, p = \overline{1, r}$, относятся контрольные объекты $w_i, i = 1, \dots, n$. При этом требуется, чтобы величина сходства признаков объектов классов $x_{p1}, x_{p2}, \dots, x_{pm_p} \in X_p, p = \overline{1, r}$ была не меньше заранее заданного числа δ . Здесь класс X_p состоит из m_p объектов $x_{p1}, \dots, x_{p(m_p)}$, и выполняется $X = \bigcup_{p=1}^r X_p$. В пространстве номинальных признаков величина, характеризующая сходство признаков объектов, обозначается через $\rho^j(w_i, x_{pq})$ и вычисляется на основе формулы (1), то есть

$$\rho_{pi}^j(w_i, x_{pq}) = \begin{cases} 1, & \text{agar } (w_i^j - x_{pq}^j) = 0; \\ 0, & \text{aks holda} \end{cases} \quad (14)$$

здесь $p = \overline{1, r}; i \neq q = \overline{1, m_p}; j = \overline{1, N}$;

Указанная величина является параметрами вектора и представляется в следующем виде $\rho_{pi}(w_i, x_{pq}) = (\rho_{pi}^1(w_i, x_{pq}), \rho_{pi}^2(w_i, x_{pq}), \dots, \rho_{pi}^N(w_i, x_{pq}))$.

Следовательно, для произвольных объектов w_i и x_{pq} p -го класса в булевом векторном пространстве определяется булев вектор

$$\rho_{pi}(w_i, x_{pq}) = (\rho_{pi}^1(w_i, x_{pq}), \rho_{pi}^2(w_i, x_{pq}), \dots, \rho_{pi}^N(w_i, x_{pq})).$$

Компоненты данного вектора характеризуют сходство значимых признаков двух объектов в разрезе рассматриваемого признака. Если $\rho_{pi}^j(w_i, x_{pq}) = 1$, то объекты w_i и x_{pq} участвуют в вычислениях по j -му признаку, в противном случае, то есть при $\rho_{pi}^j(w_i, x_{pq}) = 0$, соответствующий j й признак не участвует в вычислительных процедурах. $\Gamma_p(w_i, X_p) = \frac{1}{m_p} \sum_{k=1}^{m_p} (\rho_{pi}(w_i, x_{pk}))$ обозначает среднюю степень сходства контрольного объекта w_i с объектами класса X_p . Аналогично, величина $\Gamma_q(w_i, X_q) = \frac{1}{m_q} \sum_{k=1}^{m_q} (\rho_{qi}(w_i, x_{qk}))$ характеризует среднюю степень сходства объекта w_i со всеми объектами класса X_q . В качестве решающего правила используется алгоритм вычисления полученных оценок.

$$\mathfrak{R}(W)|_\lambda = \begin{cases} p, & \text{agar } |\Gamma_p|_\lambda - |\Gamma_q|_\lambda \geq \delta, q \neq p, q = \overline{1, r} \\ 0, & \text{boshqa barcha hollarda} \end{cases} \quad (15)$$

здесь $|\Gamma_q|_\lambda = \frac{1}{m_p} \sum_{k=1}^{m_p} (\rho_{pi}(w_i, x_{pk}))|_\lambda$ функция близости, а δ - заранее заданное некоторое положительное число (показатель надёжности). Диагностический алгоритм, ориентированный на минимизацию ошибок. Данный алгоритм направлен на классификацию объектов в пространстве номинальных признаков путём оценки средней степени сходства объектов.

Представление алгоритма диагностики, направленного на минимизацию ошибок.

шаг 1. Объекты общей обучающей выборки вводятся в базу данных.

Исходная база данных формируется по объектам всех классов $X_p, p = \overline{1, r}$.

шаг 2. Для контрольных объектов w_i на основе приведённых выше формул (1) вычисляются все параметры вектора $\rho_{pi}(w_i, x_{pk})$ для всех $p = \overline{1, r}$; $i = \overline{1, n}$; $k = \overline{1, m_p}$; $j = \overline{1, N}$.

шаг 3. Для каждого сформированного класса определяются значения величин, характеризующих среднюю степень сходства объектов $\Gamma_p(w_i, X_p), \Gamma_q(w_i, X_q)$ где $p, q = \overline{1, r}$

шаг 4. Определяется степень сходства всех контрольных объектов w_i по отношению к классам. Если выполняется условие $|\Gamma_p|_\lambda - \Gamma_q|_\lambda| \geq \delta, q \neq p, q = \overline{1, r}$, то объект считается принадлежащим классу p .

В результате добавления синтетических объектов показатели точности моделей возросли. Это свидетельствует о том, что увеличение объёма данных положительно влияет на качество кластеризации и классификации (рисунок 4).

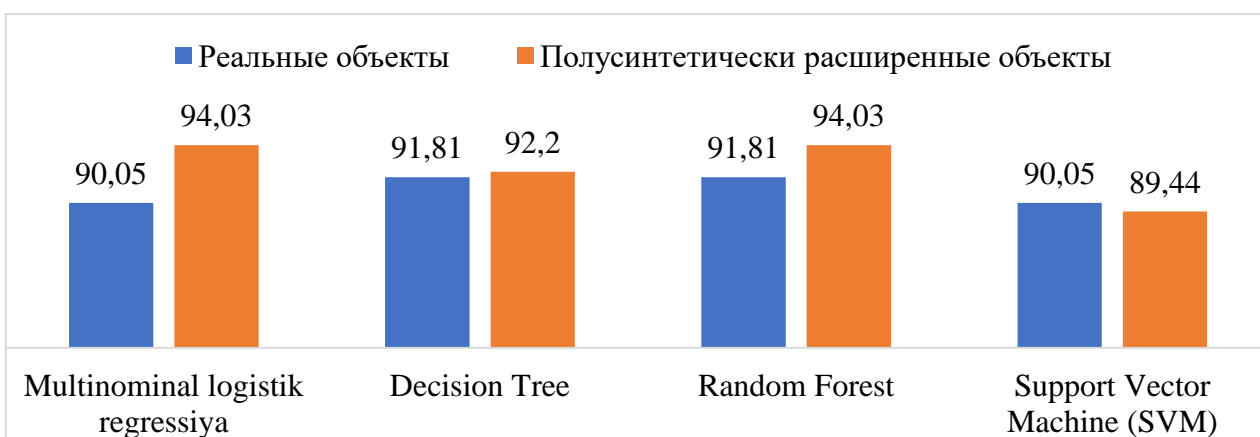


Рисунок 4. Анализ реальных и синтетических объектов

Результаты, полученные на основе синтетических объектов, а также показатели после отбора информативных признаков были сопоставлены. Полученные результаты показывают, что при применении алгоритма отбора информативных признаков во всех моделях наблюдается существенное повышение точности, что подтверждает эффективность предложенного подхода (рисунок 5).

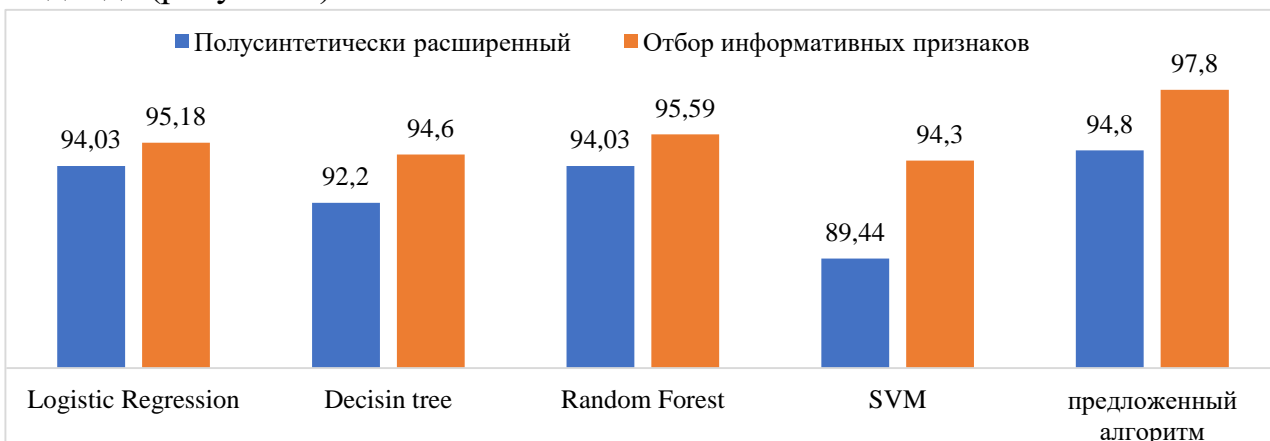


Рисунок 5. Анализ синтетических объектов и выбор информативных признаков

В четвёртой главе диссертации, озаглавленной «Применение программного комплекса SBSK для решения практических задач медицины», подробно изложены общая структура, функциональные возможности, графический интерфейс и механизм функционирования программного комплекса SBSK, созданного на основе алгоритмов, разработанных в рамках диссертационной работы.

В параграфе 4.1 рассматривается система «Smart Diagnosis», разработанная в виде веб-платформы, которая осуществляет диагностику на основе анализа медицинских признаков и анкетных данных пользователей. Процесс диагностики основан на совокупности алгоритмов машинного обучения, оптимального отбора признаков и методов минимизации ошибок классификации. Платформа обрабатывает введённые пользователем данные, формирует диагностические результаты и обеспечивает поддержку принятия медицинских решений. Практические аспекты системы, реализованные алгоритмы, точность диагностики и интеграция с веб-интерфейсом проанализированы подробно. Результаты исследования показывают, что предложенная система позволяет выявлять рак молочной железы на ранних стадиях с высокой точностью и имеет значительную практическую значимость в современной медицинской диагностике.

В параграфе 4.2 разработан программный комплекс SBSK как инновационная информационная система, обеспечивающая анализ субъективных симптоматических данных, полученных от пользователя, с применением технологий искусственного интеллекта и представление результатов пользователю через графический интерфейс. Освещены функциональные возможности программного средства, архитектурные компоненты, backend-система, взаимодействие с базой данных, а также механизм совместной работы технологий Python и PHP.

ЗАКЛЮЧЕНИЕ

По результатам исследований, проведённых в рамках диссертационной работы на тему «**Диагностический алгоритм, основанный на минимизации ошибок при раннем выявлении рака молочной железы**», сформулированы следующие выводы.

1. На основе историй болезни пациентов с раком молочной железы проведён анализ формирования набора признаков и общей выборки, методов предварительной обработки данных, а также критериев и алгоритмов оценки значимости объектов и признаков. Усовершенствован адаптивный алгоритм очистки клинических данных рака молочной железы, позволяющий автоматически устранять шум, пропуски и неопределённости, присутствующие в обучающей выборке и датасете. В результате существенно повышено качество данных и сформирована надёжная очищенная выборка для последующих этапов кластеризации и диагностики, что внесло значительный вклад в повышение общей эффективности диагностики.

2. Разработан алгоритм микрокластеризации на основе подхода минимального остовного дерева (MST) при диагностике объектов. Данный алгоритм обеспечивает выявление скрытых структурных зависимостей в

данных по раку молочной железы с высокой точностью и эффективностью даже в условиях высокой размерности. В результате созданное программное средство послужило инструментом поддержки принятия решений для врачей, способствуя оценке состояния пациентов и определению необходимости их направления к специалистам-онкологам и радиологам. Это позволило повысить производительность труда специалистов в среднем на 20 %.

3. Разработан алгоритм отбора набора информативных признаков на основе коэффициента ошибок классификации и теории вероятностей. Из 32 признаков, характеризующих рак молочной железы, отобран набор из 18 наиболее информативных признаков, используемых для диагностики. Данный подход позволил снизить сложность диагностической модели при сохранении её точности и устойчивости, а также сократить время постановки диагноза примерно на 20–25 %.

4. Разработан алгоритм диагностики, основанный на минимизации различий степеней сходства объектов классов в пространстве информативных признаков и диагностических ошибок. Это позволило радиологам и онкологам своевременно применять необходимые медицинские меры при диагностике рака молочной железы. По результатам клинических испытаний время, затрачиваемое на анализ диагностических данных, сократилось на 20 %, а эффективность диагностики повысилась в 1,7 раза.

5. На основе результатов исследования разработаны и внедрены в практику программные комплексы «SBSK» и «SmartDiagnosis». Данные решения успешно применяются в ряде медицинских учреждений и сформировали универсальную методологическую основу для анализа сложных клинических данных.

**SCIENTIFIC COUNCIL AWARDING SCIENTIFIC DEGREES
DSc.09/2025.27.12.T.01.03 AT TASHKENT UNIVERSITY OF
INFORMATION TECHNOLOGIES**

DENAU INSTITUTE OF ENTREPRENEURSHIP AND PEDAGOGY

KHAYDAROV SHERALI

**A DIAGNOSTIC ALGORITHM BASED ON MINIMIZING ERRORS IN
THE EARLY DETECTION OF BREAST CANCER**

05.01.03 – Theoretical fosis computer science

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

Tashkent - 2026

The theme of Doctor of Philosophy (PhD) dissertation in technical sciences was registered at the Supreme Attestation Commission at the Ministry of Higher Education, Science and Innovations of the Republic of Uzbekistan under number B2025.2.PhD/T5627.

The dissertation was carried out at the Denau Institute of Entrepreneurship and Pedagogy. The abstract of the dissertation is posted in three languages (Uzbek, Russian, English (resume)) on the website of the Scientific Council (www.tuit.uz) and on the "Ziyonet" Information and educational portal (www.ziyonet.uz).

Scientific adviser: Mamajanov Rakhmatilla Yakubzhanovich
Candidate of Technical Sciences, Associate Professor

Official opponents: Khudaiberdiev Mirzaakbar Khakkulmirzaevich
Doctor of Technical Sciences, Senior Researcher

Boboyev Lochinbek Boymurotovich
Doctor of Philosophy (PhD) in technical sciences

Leading organization: Fergana State Technical University

The defense of dissertation will take place on "10" June 2026 at 16⁰⁰ o'clock at the meeting of the Scientific Council No. DSc.09/2025.27.12.T.01.03 at Tashkent University of Information Technologies. (Address: 108, Amir Temur Street, Tashkent city, 100084. Tel.: (+99871) 207-59-42, e-mail: ilmiy_kengash@tuit.uz)

The dissertation is available at the Information Resource Centre of Tashkent University of Information Technologies named after Muhammad al-Khwarizmi (Registration number № 416). (Address: 108, Amir Temur Street, Tashkent city, 100084. Tel.: (+99871) 238-64-15)

The abstract of the dissertation was sent out on "22" May 2026 year.
(mailing report № 23, on "22" May 2026 year).



M.M.Kamilov
Chairman of the Scientific Council for
awarding scientific degrees, Doctor of
Technical Sciences, Academician of the ASU

N.A.Egamberdiev
Scientific Secretary of the Scientific
Council for awarding scientific degrees,
Doctor of Philosophy in Technical Sciences

N.O.Rakhimov
Chairman of the scientific seminar under the
Scientific Council for awarding scientific degrees,
Doctor of Technical Sciences, Docent

INTRODUCTION (abstract of PhD dissertation)

Objective of the research: To develop an algorithm for minimizing diagnostic errors in the early detection of breast cancer.

Object of the research. The processes of preliminary processing of medical data, clustering, and minimizing errors in object diagnosis within intelligent data analysis.

Subject of the research. Preliminary data processing, microclustering-based clustering, selection of a set of informative features, and diagnostic algorithms and software based on error minimization in object diagnosis.

Research methods. In the course of the research, methods of preliminary processing of medical data, microclustering aimed at improving reference accuracy, the error coefficient and probability theory for selecting informative features, and early diagnostic methods were used.

The scientific novelty of the research is as follows:

Scientific novelty of the research consists of the following:

An adaptive cleaning algorithm for preliminary processing, aimed at identifying and eliminating noisy, repetitive, missing, and inconsistent values encountered in clinical data on breast cancer, has been improved based on probability theory;

An algorithm for microclustering based on a minimum spanning tree, aimed at improving grouping accuracy while taking into account degrees of proximity and similarity in object diagnosis, has been developed;

An algorithm for selecting a set of informative features, aimed at evaluating the significance level of features based on the error coefficient and probability theory, as well as extracting the most important characteristics in object classification, has been developed;

A diagnostic algorithm, aimed at determining the difference in similarity degrees among class objects in the space of informative features, minimizing errors arising during the diagnostic process, and improving the accuracy of decision-making, has been developed.

Implementation of research results. Based on the results obtained from using the diagnostic algorithm developed in the study which is based on minimizing errors in the early detection of breast cancer diseases and the "Smartdiagnosis" software complex:

An adaptive cleaning algorithm for preliminary processing of clinical data on breast cancer, based on probability theory, has been developed, and the "SBSK" software has been implemented in practice at the "Surkhandarya Branch of the Republican Specialized Scientific-Practical Medical Center of Oncology and Radiology" (Termez city, Surkhandarya region, Republic of Uzbekistan) with the aim of improving diagnostic processes (Reference No. 02-28/26482 dated November 5, 2025, from the Ministry of Health of the Republic of Uzbekistan). As a result, from 32 symptomatic signs describing breast cancer, a set of 18 most significant informative features was selected for diagnosis; based on these features,

the time required for making a diagnosis was reduced by 20–25%, and diagnostic reliability was increased by 20%;

A microclustering algorithm based on a minimum spanning tree for diagnosing objects through degrees of proximity and similarity has been developed, and the “Smartdiagnosis” software has been implemented in practice at the “Termez Branch of Tashkent State Medical University” (Termez city, Surkhandarya region, Republic of Uzbekistan) with the aim of improving diagnostic processes and supporting decision-making (Reference No. 02-28/26482 dated November 5, 2025, from the Ministry of Health of the Republic of Uzbekistan). As a result, by using the selected set of informative features, the time required for diagnosing the disease was reduced by 1.5 times, and diagnostic efficiency was increased by 1.7 times;

A diagnostic algorithm aimed at reducing the error coefficient in the space of informative features for minimizing errors in object diagnosis has been developed, and the “Smartdiagnosis” software has been implemented in practice at the “Utan Polvon Diagnostic and Treatment Center” (Termez city, Surkhandarya region, Republic of Uzbekistan) with the aim of improving diagnostic efficiency (Reference No. 02-28/26482 dated November 5, 2025, from the Ministry of Health of the Republic of Uzbekistan). As a result, based on the developed algorithms, the processes of breast cancer diagnosis were automated, and a 20% increase in the reliability of the diagnostic process was confirmed.

The structure and volume of the dissertation. The dissertation consists of an introduction, four chapters, a conclusion, a list of references, and appendices. The total volume of the dissertation is 118 pages.

E'LON QILINGAN ISHLAR RO'YXATI
СПИСОК ОПУБЛИКОВАННЫХ РАБОТ
LIST OF PUBLISHED WORKS

I bo'lim (Часть I; Part I)

1. Nishanov A. Kh., Mamajanov R. Y., Xaydarov Sh. I., Mengturayev F. Z., Yuldashev R.R. Sut bezi saratoni kasalliklarini simptomlarini mashinali o'qitishga tayyorlash bosqichlari // Scientific journal of "Digital transformation and artificial intelligence" Vol. 2, Issue 6, 2024. -P.237-249 (05.00.00 04.07.2023. 340/5-son rayosat qarori).

2. Nishanov A. Kh., Mamajanov R. Y., Xaydarov Sh. I., Mengturayev F. Z. Saraton kasalliklarini erta aniqlashning muhimligi va zamonavaiy texnologiyalarga asoslangan usul va algoritmlari // Scientific journal of "Digital transformation and artificial intelligence" Vol. 3, Issue 1, 2025. -P.110-117 (05.00.00 04.07.2023. 340/5-son rayosat qarori).

3. Хайдаров Ш. И. Критерии, алгоритмы и передовые подходы, используемые в диагностике рака // Научно-технический журнал "Информационные технологии моделирования и управления". №1(139). "Научная книга". -С. 11-21 (05.00.00; №43).

4. Nishanov A. Kh., Mamajanov R. Y., Xaydarov Sh. I., Mengturayev F. Z., Yuldashev R.R. Sut bezi saratoni diagnostikasining matematik modellari // "Al-Farg'oniy avlodlari" jurnali Vol. 1, Issue 1, 2025. -P.6-14 (05.00.00 30.09.2023. 343-sonli rayosat qarori).

5. Хайдаров, Ш. И. Постановка задачи оптимизации, основанной на минимизации ошибок диагностики // Современные инновации, системы и технологии-Modern Innovations, Systems and Technologies, Vol. 1, Issue 1, 2025. -P.2010-2016. (№ 25 Directory of Open Access Journals)

6. Nishanov A.X., Xaydarov Sh.I., Mamajanov R.Ya., Yuldashev R.R. Sut bezi saratoni kasalligini erta tashxislashga qaratilgan sintetik o'quv tanlanmani yaratish algoritmi // "Muhammad al-Xorazmiy avlodlari" ilmiy-amaliy va axborot-tahliliy jurnali. №1(31). Toshkent-2025. -B. 67-74 (05.00.00 №10).

7. Хайдаров Ш.И. Алгоритмы вычисления оценок, основанные на минимизации ошибок при классификации объектов // Современные инновации, системы и технологии-Modern Innovations, Systems and Technologies. Vol. 5, Issue 2, 2025. -P. 5021-5028. (№ 25 Directory of Open Access Journals).

8. Xaydarov Sh. I. Sut bezi saratonini erta aniqlash uchun sun'iy intellekt va optimallashtirish algoritmlariga asoslangan veb -platformani // "Al-Farg'oniy avlodlari" jurnali Vol. 1, Issue 2, 2025. -P.7-14. (05.00.00 30.09.2023. 343-sonli rayosat qarori).

9. Nishanov A. X., Mamajanov R.Y., Xaidarov, Sh. I. Sut bezi saraton kasalliklarini erta aniqlashda informativ belgilar majmuasini tanlash algoritmi. // Muhandislik va iqtisodiyot jurnali Vol. 2, Issue 5, 2024. -P.6-14 (05.00.00 28.08.2024 360/5-sonli rayosat qarori).

II bo‘lim (Часть II; Part II)

10. Nishanov A.X., Xaydarov Sh.I., Mamajanov R.Ya., Xolbekov A. M., Karimova Z. B. Diagnostic Algorithm for Early Detection of Breast Cancer Based on Error Minimization Approach // International Journal of Innovative Science and Research Technology (IJISRT). Volume 9, Issue 12, 2024 –P.1535-1542. (№23 Scientific Journal Impact Factor, SJIF=8.716).

11. Nishanov A. Kh., Mamajanov R. Y., Xaydarov Sh. I., Yuldashev R.R. Sut bezi saratoni simptomlarini mashinali o‘qitishga tayyorlash // “Oliy ta’lim tizimini raqamlashtirishning dolzarb muammo va yechimlari” mavzusidagi xalqaro ilmiy-amaliy konferensiyasi ma’ruzalar to‘plami, Jizzax -2025. -B.1080-1084.

12. Xaydarov Sh. I. Sut bezi saratoni diagnostikasining matematik qo‘yilishi // “Sun’iy intellekt va kompyuter texnologiyalari: amaliyot, tajriba, muammo va istiqbollari” mavzusidagi xalqaro ilmiy-amaliy anjuman materiallari Denov – 2025. –B.310–319.

13. Nishanov A. X., Mamajhanov, R. Y., Xaydarov Sh. I. Saraton kasalliklarini erta aniqlashning ahamiyati va zamonaviy texnologiyalar asosidagi metod va algoritmlar // “Sun’iy intellekt va kompyuter texnologiyalari: amaliyot, tajriba, muammo va istiqbollari” mavzusidagi xalqaro ilmiy-amaliy anjuman materiallari Denov – 2025. -B.303-310.

14. Нишанов А. Х., Мамаджанов, Р. Я., Хайдаров, Ш. И. Критерии, алгоритмы и передовые подходы, используемые в диагностике рака // Сборник материалов Республиканской научно-практической конференции на тему «Современные проблемы интеллектуальных систем» Часть 1, -С. 250-256 Джизаке. 2025.

15. Xaydarov Sh. I. Sun’iy intellekt optimallashtirish asosida sut bezi saratonini erta aniqlash platformasi // “Kompyuter ilmlari va muhandislik texnologiyalari” mavzusidagi xalqaro ilmiy-texnik anjuman materiallari (1-qism. -B. 28–34). Jizzax -2025

16. Xaydarov Sh. I., Chorshanbiyeva M. Clustering algorithms for breast cancer detection // “Sun’iy intellekt va kompyuter texnologiyalari: amaliyot, tajriba, muammo va istiqbollari” xalqaro ilmiy-amaliy anjuman materiallari 1-qism. Denov-2025. -B. 332-339.

17. Nishanov A.X., Mamajanov R.Y., Xaydarov, Sh. I. Obyektlarni sinflashtirishda xatoliklarni minimallashtirishga asoslangan baholarni hisoblash algoritmlari. // “Ta’limning zamonaviy usullari va innovatsion texnologiyalar” xalqaro konferentsiya to‘plami. Tashkent 2025. www.mmit.tiue.uz DOI: 10.61587/mmit.tiue.uz.v1i1.176 -B. 151-157.

18. Xaydarov Sh. I., Nishanov A.X., Mamajanov R.Ya., Turdiyeva S.I., Elmurodova T.L., Shaximardonova X.B. Sut bezi saratonini erta tashxislashda sintetik tanlama yaratishning amaliy dasturi // O‘zbekiston Respublikasi Adliya vazirligi. EHM uchun yaratilgan dasturning rasmiy ro‘yxatdan o‘tkazilganligi to‘g‘risidagi guvohnoma №DGU 48327. 05.03.2025.

19. Nishanov A.X., Mamajanov R.Ya., Xaydarov Sh. I., Urozov I.F., Isoqov F.F. Sut bezi saratoni kasalliklarini erta aniqlashga ko‘mak beruvchi dasturiy majmua

ma'lumotlar bazasi // O'zbekiston Respublikasi Adliya vazirligi. Ma'lumotlar basining rasmiy ro'yxatdan o'tkazilganligi to'g'risidagi guvohnoma №BGU 1821. 11.02.2025.

20. Xaydarov Sh. I., Nishanov A.X, Yuldashev R.R. Sut bezi saratoni kasalliklarini erta aniqlashda xatoliklarni minimallashtirishga asoslangan tashxislash algoritmi // O'zbekiston Respublikasi Adliya vazirligi. EHM uchun yaratilgan dasturning rasmiy ro'yxatdan o'tkazilganligi to'g'risidagi guvohnomasi №DGU 45422. 12.12.2024.

Avtoreferat «Texnika fanlari va innovatsiyalar» jurnali tahririyatida tahrirdan o‘tkazilib, o‘zbek, rus va ingliz tillaridagi matnlar o‘zaro muvofiqlashtirildi.

Bosmaxona litsenziyasi:



9338

Bichimi: 84x60 ¹/₁₆. «Times New Roman» garniturası.
Raqamli bosma usulda bosildi.
Shartli bosma tabog‘i: 4. Adadi 100 dona. Buyurtma № 31/25.

Guvohnoma № 851684.
«Tipograff» MCHJ bosmaxonasida chop etilgan.
Bosmaxona manzili: 100011, Toshkent sh., Beruniy ko‘chasi, 83-uy.