МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

СУМСЬКИЙ ДЕРЖАВНИЙ УНІВЕРСИТЕТ

КАФЕДРА КОМП'ЮТЕРНИХ НАУК

КВАЛІФІКАЦІЙНА МАГІСТЕРСЬКА РОБОТА

на тему:

«Моделі та метод інформаційної технології діагностування інфекційних хвороб»

Завідувач випускаючої кафедри — **Довбиш А.С.**

Керівник роботи — **Довбиш А.С.**

Студент групи ІН.мз-92с — **Шуда І.О.**

СУМИ 2021

Сумський державний університет
(назва вузу)

Факультет Еліт Кафедра Комп'ютерних наук
Спеціальність «Комп'ютерні науки»

Затверджую:
зав.кафедри _____
_____
"_____" _____ 20___р.

ЗАВДАННЯ
НА ДИПЛОМНИЙ ПРОЕКТ (РОБОТУ) СТУДЕНТОВІ

Шуді Ірині Олександрівні
(прізвище, ім'я, по батькові)

1. Тема проекту (роботи)
Моделі та метод інформаційної технології діагностування інфекційних хвороб

затверджено наказом по інституту від "____" _____ 20 ___ р. № _____

2. Термін здачі студентом закінченого проекту (роботи)
26.01.2021 р.

3. Вхідні данні до проекту (роботи)
Архівні дані результатів лабораторно-клінічних аналізів, надані Сумською клінічною лікарнею інфекційних хвороб ім. Красовицького

4. Зміст розрахунково-пояснювальної записки (перелік питань, що їх належить розробити)
1) Аналіз проблеми дослідження
2) Опис методів досліджень
3) Інформаційне, алгоритмічне та програмне забезпечення системи діагностування інфекційних хвороб.
4) Додаток А. Лістінг програми.

5. Перелік демонстраційного матеріалу
1)Актуальність; 2) Мета; 3) Постановка задачі; 4) Категорійні моделі; 5) Критерій оптимізації параметрів машинного навчання; 6) Схема алгоритму машинного навчання; 7) Результати моделювання; 8) Висновки

6. Консультанти до проекту (роботи), із значенням розділів проекту, що стосується їх

Розділ	Консультант	Підпис, дата	
		Завдання видав	Завдання прийняв

7.Дата видачі завдання_____

Керівник _____
(підпис)

Завдання прийняв до виконання _____
(підпис)

КАЛЕНДАРНИЙ ПЛАН

№ п/п	Назва етапів дипломного проекту (роботи)	Термін виконання проекту (роботи)	Примітка
1.	*Огляд літератури згідно теми диплома*		
2.	*Аналіз проблеми дослідження*		
3.	*Опис методу розв'язання поставленої задачі.*		
4.	*Інформаційний синтез системи діагностування інфекційних хвороб*		
5.	*Програмна реалізація алгоритму машинного навчання*		
5.	*Оформлення пояснювальної записки до дипломної роботи*		

Студент – дипломник _____
(підпис)

Керівник проекту _____
(підпис)

РЕФЕРАТ

Записка: 59 стор., 8 рис., 1 табл., 20 джерел інформації.

Мета роботи – підвищення функціональної ефективності машинного навчання системи діагностування патологічних процесів.

Об'єкт дослідження – процес машинного навчання системи діагностування патологічних процесів.

Предмет дослідження – категорійні моделі, критерій та метод оптимізації параметрів машинного навчання, вирішальні правила та засоби інформаційної технології діагностування патологічних процесів.

Метод дослідження – теоретико-інформаційний підхід до оцінки інформаційної спроможності системи діагностування патологічних процесів, багатовимірний статистичний аналіз для визначення базового класу розпізнавання та інформаційно-екстремальний метод машинного навчання системи діагностування патологічних процесів.

Розроблено алгоритмічне та програмне забезпечення системи діагностування патологічних процесів в рамках інформаційно-екстремальної інтелектуальної технології аналізу даних, яка базується на максимізації інформаційної спроможності системи в процесі машинного навчання. Досліджено вплив параметрів машинного навчання на функціональну ефективність системи діагностування патологічних процесів за результатами клініко-лабораторних досліджень.

ІНФОРМАЦІЙНО-ЕКСТРЕМАЛЬНА ІНТЕЛЕКТУАЛЬНА ТЕХНОЛОГІЯ, МАШИННЕ НАВЧАННЯ, ОПТИМІЗАЦІЯ, ІНФЕКЦІЙНА ПАТОЛОГІЯ, ІНФОРМАЦІЙНИЙ КРИТЕРІЙ, ВИРІШАЛЬНЕ ПРАВИЛО,

ЗМІСТ

ВСТУП

За останні роки різко зріс вплив комп’ютерних та інформаційних технологій на всі сфери діяльності людства. Особливо стрімкого розвитку вони набули з виникненням глобальної загрози людству через поширення пандемії COVID-19. Виникла потреба мінімізувати контакти людей, але при цьому забезпечити їм достойну якість життя, задовольнити їх потреби, продовжити функціонування навчання, науки, економіки, промисловості, оборони, можливості лікуватись, одружуватись та народжувати дітей. За умови пандемії важливого значення набуває здатність медичних закладів здійснювати для вибору правильної схеми лікування швидкий та достовірний експрес-діагноз патологічних процесів. Використання новітніх методів діагностування та сучасного лікувально-діагностичного обладнання дозволяє лікарю отримувати великий обсяг діагностичної інформації. Але оскільки остаточний діагноз приймає лікар, який несе за його достовірність юридичну відповідальність, то для успішного аналізу великих обсягів даних, які включають як символьні образи, так і медичні зображення, вирішальне значення мають професіональний рівень лікаря та його практичний досвід. Тому створення комп’ютеризованих систем діагностування патологічних процесів, які відносяться до класу інтелектуальних систем підтримки прийняття рішень, є актуальною задачею, розв’язання якої дозволить підвищити достовірність та оперативність прийняття діагностичних рішень.

В магістерській кваліфікаційній роботі в рамках вітчизняної так званої інформаційно-екстремальної інтелектуальної технології (IEI-технології) аналізу даних, яка ґрунтується на максимізації інформаційної спроможності системи розпізнавання в процесі машинного навчання, розроблено вхідний математичний опис, алгоритмічне та програмне забезпечення здатної навчатися системи діагностування інфекційної патології.

1 АНАЛІЗ ПРОБЛЕМИ ТА ПОСТАНОВКА ЗАДАЧІ ДОСЛІДЖЕНЬ

1.1 Сучасний стан та тенденції розвитку комп'ютеризованих систем діагностування патологічних процесів

В теперішній час розвиток медичної галузі характеризується появою нових методів діагностування, лікування й прогнозування перебігу та наслідків патологічних процесів. При цьому збільшення кількості інформації, яка надходить від засобів сучасного діагностично-лікувального обладнання, відбувається за незмінної інформаційної спроможності лікаря, що впливає на його здатність до сприйняття та осмислення даних і призводить до помилкових рішень під час вибору та проведення лікувально-профілактичних заходів. Тому через суттєве збільшення обсягу діагностичної інформації функціональна ефективність процесів діагностування, лікування та реабілітації пацієнтів все ще суттєво залежить від професійного рівня та досвіду лікаря. Цей факт обумовив необхідність інтенсивного впровадження в закладах практичної медицини інтелектуальних інформаційних технології аналізу даних через такі фактори [**Ошибка! Источник ссылки не найден.**]:

- обмеженість часових ресурсів;
- складність залучення достатнього числа висококваліфікованих експертів відповідної предметної області;
- неповна інформація про функціональний стан пацієнта;
- довільні початкові умови патологічного процесу;
- велика кількість діагностичних ознак розпізнавання захворювань;
- суттєвий перетин класів розпізнавання, що характеризують функціональні стани патологічних процесів;
- нестаціонарність, імплементивність і висока динаміка перебігу патологічних процесів.

Технологічним інструментом для визначення і планування усіх ресурсів медичного закладу, які необхідні для ведення лікувально-діагностичної, адміністративно-господарської, фінансової, сервісної діяльності та обліку в процесі надання послуг є медичні інформаційні системи (МІС). Прикладами таких МІС є «Доктор Елекс», «Медик», «ЕКСіМЕД», «TherDep», «Медістар», «Інтерін», «Артеміда», «Амулет», «Indivo» та інші, що дозволяють автоматизувати всі аспекти діяльності медичної установи та зберігати інформацію в електронній карті пацієнта (ЕКП) [2, 3]. Такі системи забезпечують інтеграцію ЕКП з різноманітним діагностичним обладнанням, і дозволяють отримувати дані безпосередньо з лабораторних аналізаторів. Подальшим технологічним рішенням у цій сфері є розроблення та впровадження єдиного інформаційного простору (порталів) для ведення ЕКП, що забезпечує інтеграцію медичних даних з різних МІС.

Важливою умовою розбудови як єдиного інформаційного простору в цілому є розробка та впровадження міжнародних галузевих стандартів обміну медичними даними, наприклад, HL7 (Health Level 7) обміну, управління та інтеграції медичної документації, HL7 CDA (Clinical Document Architecture) архітектури клінічних документів, openEHR (відкрита електронна медична картка) розроблення електронних медичних документів, DICOM (Digital Imaging and Communications in Medicine) обробки, зберігання і передавання медичних зображень та інші [4]. При цьому зберігання файлів медичних зображень у DICOM-форматі та їхнє передавання комп'ютерною мережею забезпечують спеціалізовані системи архівування зображень та комунікацій (PACS-системи).

В працях [4, 5] сформовано принципи та критерії, яким повинні відповідати комп'ютеризовані системи діагностування, а саме:

- принцип достовірності та внутрішньої несуперечливості одержаних про пацієнта даних;
- принцип стійкості до великої кількості нозологічних (патологічних) форм захворювання;

- принцип інтерпретації даних на основі медичних знань;
- принцип недостатності знань про окремий випадок захворювання;
- принцип пояснення й обґрунтування прийнятого рішення;
- принцип переоцінки рішень у разі повторного проведення дослідження.

При цьому діагностичні системи повинні виконувати такі функції [5]:

- диференціальне діагностування і вибір схеми лікування серед широкого кола нозологічних форм захворювання;
- забезпечення достовірності діагностичних рішень незалежно від ступеня клінічного прояву патологічного процесу;
- аналіз динаміки перебігу патологічного процесу;
- оцінка поточного стану пацієнта.

Початковий етап у запровадженні інформаційних інтелектуальних технологій в галузі охорони здоров'я пов'язаний із розробкою і впровадженням експертних систем (ЕС), здатними оперувати знаннями в певній предметній області з метою підтримки прийняття рішення [5,6]. У процесі навчання ЕС з залученням експертів предметної області формується база знань, яка містить формалізовані знання першого (факти, дані та закономірності) та другого (методи подання і виведення знань) роду про предметну область задачі, що розв'язується. Вирішальне правило будується шляхом застосування до бази знань системи логічного виводу, яка, формуючи послідовність правил, моделює механізм прийняття рішення експертом предметної області.

Проектування експертних систем характеризується застосуванням концепції мінімізації області прийняття рішення. У сфері інфекційних патологічних процесів експертні системи розробляються з метою контролю результатів мікробіологічних досліджень та раннього виявлення випадків інфекційних і епідеміологічних захворювань. При цьому відомі такі експертні системи [6]: ЕС «AIDS», ЕС «HIVPCES», ЕС «CTSHIV» – для

діагностування ВІЛ та надання рекомендацій щодо вибору відповідної схеми лікування, ЕС «HEPAXPERT-I» – для інтерпретації аналізу серологічних ознак гепатиту А та В, ЕС «WHONET», ЕС «MERCURIO» – для інтерпретації результатів мікробіологічних лабораторних тестів, ЕС «GUIDON», ЕС «GermWatcher» – для виявлення ризику розвитку та поширення внутрішньолікарняних інфекцій.

Розроблення та практичне застосування ЕС обмежують такі їх основні недоліки:

- відсутність властивості адаптивності, обумовленої довільними початковими умовами патологічних процесів;
- негнучкість, яка обумовлена залежністю функціональної ефективності від повноти побудованої бази знань;
- необхідність залучення до створення бази знань висококваліфікованих фахівців-експертів;
- евристичний характер медичних знань.

Потужний розвиток інформаційних технологій сприяв виникненню нового класу автоматизованих систем – системи підтримки прийняття рішень (СППР), які, на відміну від традиційних ЕС, здатні самостійно формувати базу знань без залучення фахівця предметної області.

Основною складовою сучасних комп’ютеризованих систем діагностування є СППР для діагностування і прогнозування перебігу та наслідків патологічних процесів. При цьому основним завданням СППР є інтелектуальний аналіз вхідних даних про перебіг патологічного процесу, подаючи його результат у вигляді розбиття простору діагностичних ознак на класи розпізнавання, що характеризують можливі функціональні стани патологічного процесу [1].

Одним із перспективних застосувань СППР є її включення до структури GRID-системи, яка представляє собою глобальну географічно розподілену інфраструктуру інформаційно-комунікаційних технологій для скоординованого, гнучкого та захищеного розподілу обчислювальних

ресурсів і ресурсів для накопичення та зберігання інформації. Концепцію GRID-систем у галузі охорони здоров'я можливо використовувати у таких аспектах [7]:

- забезпечення доступу до медичних даних та первинних результатів досліджень пацієнта при побудові єдиного інтегрованого інформаційного середовища системи охорони здоров'я (e-Health) для розв'язання задач діагностування та прогнозування перебігу та наслідків патологічного процесу;
- проведення епідеміологічного аналізу, тобто використання накопичених медичних даних для пошуку залежності між даними, факторами ризику, симптомами, захворюваннями з використанням методів інтелектуального аналізу даних.

Завданням GRID-системи для діагностування медичних графічних зображень є:

1) формування медичного графічного зображення за допомогою діагностичного обладнання;

2) оброблення та передача через термінал медичного графічного зображення в ресурсний центр GRID-системи із застосуванням стиснення та завадо захищеного кодування відеоінформації, що передається;

3) розпізнавання за допомогою інтелектуальної СППР ресурсного центру GRID-системи медичного графічного зображення (3D-моделювання органів; епідеміологічні дослідження, у тому числі ідентифікація генів складних захворювань, аналіз несприйняття бактерій до антибіотиків, аналіз дії ліків);

4) видання на термінал лікаря-користувача діагностичного рішення, яке має рекомендаційний характер.

Схематично структуру GRID-системи представлена на рис. 1.1 [8]. Аналіз рис. 1.1 показує, що основними структурними компонентами GRID-системи є:

1) сукупність персональних комп'ютерів зі встановленими інтерфейсами користувача для забезпечення доступу користувача до ресурсів GRID-системи;

2) сукупність ресурсних центрів, що включають в себе надавані обчислювальними елементами обчислювальні ресурси, що виконують обробку даних, і ресурси зберігання даних;

3) сукупність базових GRID-сервісів. До їх числа входить система управління даними, функцією якої є забезпечення доступу до систем зберігання даних, що існують у ресурсних центрах.

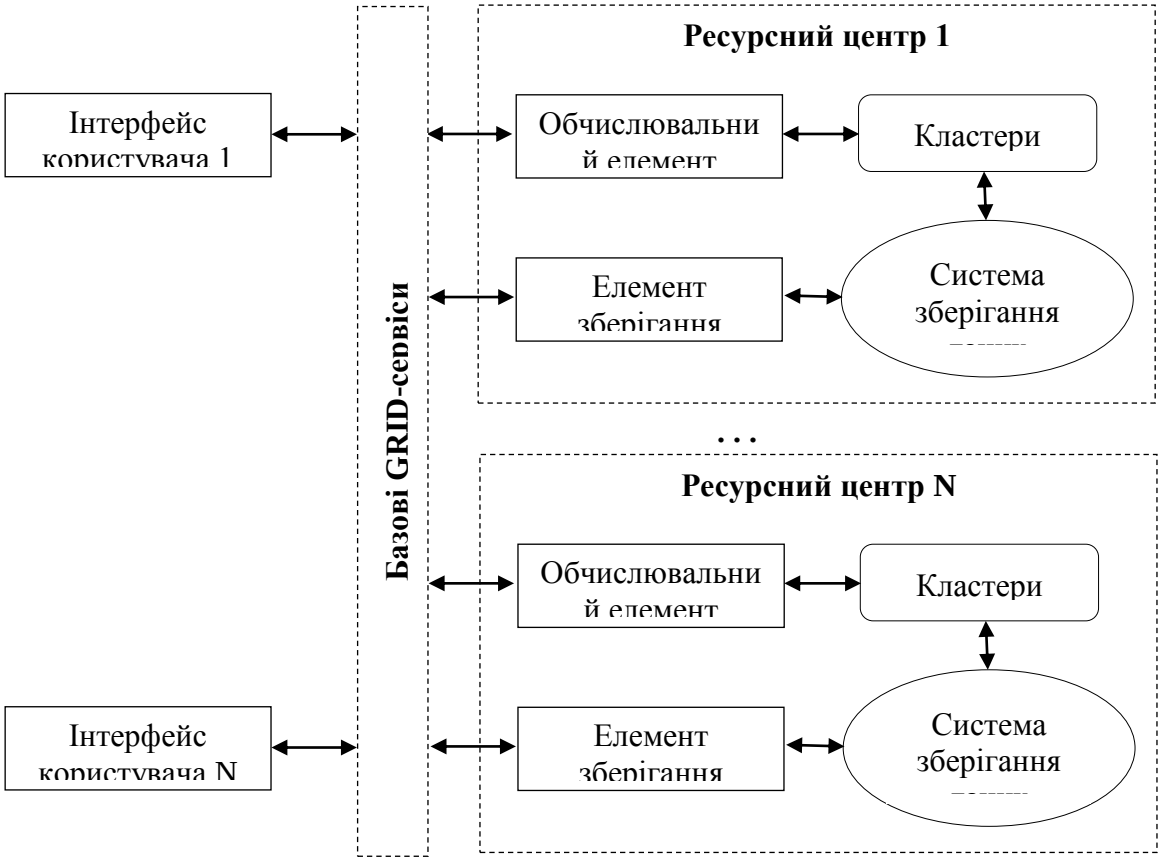


Рисунок 1.1 – Структура GRID-системи

З метою побудови СППР для діагностування патологічних процесів застосовано традиційні інформаційні технології, які реалізують моделі статистичного та регресійного аналізу, дискримінантного аналізу, моделі Маркова та інше [Вас]. Але оскільки діагностична СППР функціонує за умов

апріорної невизначеності, обумовленої довільними початковими умовами процесів діагностування і лікування та дії неконтрольованих збурюючих факторів, то автори праці [1] вважають, що підвищення функціональної ефективності СППР пов’язано з наданням їй властивості адаптивності шляхом використання ідей і методів машинного навчання та розпізнавання образів.

1.2 Аналіз методів машинного навчання

Система охорони здоров’я накопичила великий обсяг медичної інформації, що дозволяє створювати комплексні діагностичні системи на основі інтелектуального аналізу даних (Data Mining) [9, 10], в основі якого знаходяться технології штучного інтелекту та візуального представлення інформації. Методи інтелектуального аналізу даних дозволяють працювати зі слабо структурованими задачами, до яких можна віднести задачі медичного діагностування, що характеризуються навчальними вибірками малих розмірів, високою розмірністю простору діагностичних ознак розпізнавання та довільними початковими умовами формування векторів вхідних даних. Застосування технології інтелектуального аналізу даних в медичній галузі характеризується зростанням уваги до описових методів, які дозволяють виявляти приховані та попередньо невідомі знання, що можуть бути використані у процесі прийняття діагностичних рішень. З цією метою значна увага приділяється розв’язанню задач розбиття простору діагностичних ознак розпізнавання на класи еквівалентності, виділення інформативних ознак розпізнавання під час опрацювання медичних сигналів та зображень. Також цей підхід можливо використовувати у сферах медицини, де знання обмежені і лікарю, що приймає рішення, не вдається зробити висновки про результат перебігу патологічного процесу, наприклад, для визначення ефективності застосування лікарських засобів при виявленні

нових штамів патогенних мікроорганізмів, прогнозування вірусологічної відповіді на лікувальну терапію ВІЛ.

Одним із шляхів аналізу і синтезу СППР є використання методів розпізнавання, що базуються на нечіткому логічному виводі і описуються за допомогою лінгвістичних змінних і нечітких множин. Серед інфекційних патологій метод нечіткого логічного виводу використано для діагностування гострої респіраторної вірусної інфекції [10].

Широкого практичного використання набули методи аналізу і синтезу СППР з використанням штучних нейронних мереж, які здатні приймати рішення на основі виявлених прихованих закономірностях у багатовимірних даних [11 – 13] . При цьому широкий спектр пропонованих СППР характеризується загальною рисою відсутності єдиного універсального підходу до вибору архітектури та алгоритму навчання нейронної мережі. Тому сучасним підходом до побудови СППР з підвищеною функціональною ефективністю є застосування гібридних моделей нейронних мереж, традиційні алгоритми навчання яких доповнюють використанням нечіткої логіки, генетичних алгоритмів, алгоритмів кластерного аналізу, популяційних методів ройового інтелекту. Під час вирішення проблеми медичного діагностування найбільш широко нейронні мережі застосовуються для розв’язання задачі розпізнавання як медичних зображень, так і символьних образів.

При діагностуванні інфекційних патологій нейронні мережі використовують для діагностування вірусного гепатиту С, патологічних станів печінки у хворих на гепатит В та С. З допомогою нейронних мереж проводять диференціальне діагностування туберкульозу легень. З використанням нейронних мереж прогнозують ризик розвитку фіброзу печінки у хворих на ВІЛ з ускладненням вірусного гепатиту С та визначають стадії ВІЛ/СНІД. Нейронні мережі застосовують для діагностування сальмонельозу.

Здатні навчатися системи діагностування можуть функціонувати в таких режимах [19]:

- навчання за апріорно класифікованими навчальними матрицями (навчання з «учителем»);
- навчання за апріорно некласифікованими навчальними матрицями (самонавчання);

Найбільш поширеними алгоритмами навчання з «учителем» є штучні нейронні мережі, які в рамках біонічного підходу моделюють механізм прийняття рішень людиною [12, 13]. Найбільш відомим є метод зворотного поширення помилки (Backpropagation), який реалізує ітеративний градієнтний алгоритм з метою мінімізації помилки роботи багатошарового перцептрону та отримання бажаного виходу. Незважаючи на широке застосування цього алгоритму, він не є ідеальним. Найбільше неприємностей приносить невизначено довгий процес навчання. У складних завданнях для навчання мережі можуть знадобитися дні або навіть тижні, і вона може і взагалі не навчитися. Причиною цього може бути:

а) параліч мережі – завмирання процесу навчання як наслідок дуже малого значення похідної стискаючої функції;

б) потрапляння в локальні мінімуми;

в) неправильне визначення розміру кроку.

При цьому оскільки штучні нейронні мережі при зміні потужності словника змінюють свою структуру, то вони є чутливими до збільшення багатовимірності як словника ознак, так і алфавіту класів розпізнавання.

Використання статистичних байєсівських класифікаторів дало змогу розробити ряд алгоритмів, які використовуються у багатьох сучасних системах для оцінки простору станів деяких змінних. Найбільш відомими є наївний байєсівський класифікатор (Naive Bayes classifier) та Байєсова мережа довіри (Bayesian network) [14]. Наївний байєсівський класифікатор базується на використанні теореми Байєса з простими припущеннями про незалежність. Для оцінки параметрів моделей використовують метод

максимальної подібності, тобто, можна працювати з наївною байєсівською моделлю, не використовуючи байєсівську ймовірність та методи теорії перевірки статистичних гіпотез. Перевагою такого класифікатора є відносно невелика кількість даних для навчання.

Для підвищення точності класифікатор Байєса ціною невеликого збільшення обсягу обчислень був розроблений метод AODE (Averaged One-Dependence Estimators). Він виконує класифікацію шляхом усереднення передбачень кількох класифікаторів, у яких всі атрибути залежать від одного спільного атрибуту (класу). AODE не виконує вибір моделі і не використовує регульованих параметрів, тому як результат, він має низьку дисперсію. Він підтримує додаткове навчання класифікатора, яке здійснюється за рахунок появи нових реалізацій. При цьому прогнозує ряд ймовірностей віднесення до класів, а не просто видає степінь належності до одного конкретного класу, що дозволяє користувачеві визначити достовірність, з якою кожна класифікація може бути виконана.

Прикладом еволюційних алгоритмів є GEP (Gene expression programming) [15]. Цей алгоритм створює комп'ютерні програми, які є складними деревовидними структурами, здатними навчатися і пристосовуватися, змінюючи свої розміри, форму і склад так само, як і живий організм. Як і живі організми, комп'ютерні програми GEP також кодуються в прості лінійні хромосоми фіксованої довжини. Таким чином, це система типу генотип-фенотип, яка бере найкраще з простого геному, зберігаючи і передаючи генетичну інформацію і складний фенотип для вивчення навколишнього середовища та адаптування до нього.

Одними із ефективних методів розпізнавання образів є радіально-базисні методи, до яких слід віднести метод найближчих сусідів (Nearest Neighbor Algorithm) і його узагальнення – це метод *K*-найближчих сусідів [16]. Це сімейство алгоритмів навчання, в яких замість виконання явного узагальнення, порівнюються нові об’єкти з тими, які вже знаходяться в навчальній матриці.

Як популярний метод машинного навчання та розпізнавання образів слід вказати на метод опорних векторів (Support vector machines), який належить до групи граничних методів і визначає класи за допомогою меж просторів [17]. Опорними векторами вважаються об’єкти множини, що лежать на цих межах. Класифікація вважається вдалою, якщо простір між межами є порожнім

До класу СППР, які здатні самонавчатися (Unsupervised learning), відносять системи, що функціонують в режимах [19]:

- кластер-аналізу, який автоматизує формування вхідної навчальної матриці шляхом розбиття простору ознак на класи розпізнавання;
- факторного кластер-аналізу, задачею якого є виділення в процесі експлуатації СППР нових класів розпізнавання для донавчання;
- оптимізації словника ознак розпізнавання;
- автоматичного формування вхідного математичного опису СППР шляхом оптимізації параметрів оброблення, стиснення, квантування та фільтрації вхідної інформації.

Кластеризація або автоматична класифікація є методом неконтрольованого навчання і статистичного аналізу даних [18]. Різні методи кластеризації роблять різні припущення про структуру даних, часто визначається деяка подібність метрики і оцінювання, наприклад, внутрішні компактності (подібність між представниками одного й того ж кластеру) і поділу між різними кластерами. Інші методи засновані на оцінках щільності та графіках підключення. Кластеризація є методом неконтрольованого навчання і поширений метод для статистичного аналізу даних. Існує багато різновидностей кластеризації, які важко класифікувати, оскільки багато з них використовують одні і ті ж методи, але можна виділити наступні принципи побудови правил: ймовірнісний підхід (K-середніх (K-means), K-medians, EM-алгоритм (Expectation-maximization algorithm), алгоритми сімейства FOREL, дискримінантний аналіз); підхід на основі штучного інтелекту (нечітка кластеризація C-середніх (C-means), нейрона мережа Кохонена,

генетичний алгоритм); логічний підхід; теоретико-графовий підхід; ієрархічна кластеризація та інші.

Навчання з закріпленням (англ. Reinforcement Learning): алгоритм навчається за допомогою тактики нагороди та покарання для максимізації вигоди для агентів (систем до яких належить компонента, що навчається) Досить часто для моделювання таких алгоритмів використовується метод Монте Карло.

- Temporal difference learning або метод часових різниць є методом прогнозування, суть якого полягає в певній кореляції майбутніх прогнозів. Основна ідея полягає в тому, що коригуються прогнози, щоб відповідати іншим, більш точним, передбаченням про майбутнє.

- Q-learning – метод, застосовуваний у штучному інтелекті при застосуванні агентів. На основі одержуваної від середовища винагороди, агент формує функцію корисності Q, що згодом дає йому можливість уже не випадково вибирати стратегію поведінки, а враховувати досвід попередньої взаємодії з середовищем. Одна з переваг – те, що є можливість оцінити очікувану корисність доступних дій, не формуючи моделі навколишнього середовища. Застосовується для ситуацій, які можна представити у вигляді марковского процесу прийняття рішень. Одним із різновидів є алгоритм SARSA [160].

- Learning Automata метод реалізується як блок адаптивного прийняття рішень, розташований у довільному середовищі, який навчається через повторювані взаємодії з навколишнім середовищем. Дії вибираються відповідно з певним розподілом ймовірностей, який оновлюється залежно від реакції автомата на середовище після виконання певних дій.

Класифікаційне управління фокусується на прогнозі, засноване на відомих властивостях навчальних даних. Деякі системи намагаються усунути необхідність в людській інтуїції при аналізі даних, а інші обирають спільний підхід між людиною і машиною. Людська інтуїція не може бути повністю виключена, оскільки розробнику системи необхідно вказати, які дані повинні

бути представлені і які механізми будуть використовуватися для пошуку характеристик даних.

Нові методи контролю виникають внаслідок розроблення нових моделей інтелектуальної поведінки, а обчислювальні методи створюються для їх підтримки.

Однією із новітніх і перспективних технологій інтелектуального аналізу даних є так звана інформаційно-екстремальна інтелектуальна технологія (ІЕІ-технологія) аналізу даних, яка ґрунтується на максимізації інформаційної спроможності СППР в процесі її навчання (самонавчання) [1, 19]. ІЕІ-технологія розроблена в Сумському державному університеті і знайшла широке практичне застосування в різних галузях соціально-економічної сфери українського суспільства. Вона спрямована на усунення недоліків відомих методів класифікації та має наступні переваги:

- для прийняття рішень ІЕІ-технології використовує детерміновано-статистичний підхід, що дозволяє будувати прості детерміновані вирішальні правила, статистична корекція яких здійснюється в процесі інформаційно-екстремального машинного навчання;
- в основі даної технології закладена пряма оцінка інформаційної спроможності СППР, що навчається, шляхом використання інформаційного критерію оптимізації параметрів машинного навчання;
- методи інформаційно-екстремального машинного навчання на відміну від нейроподібних структур реалізуються в рамках функціонального підходу до моделювання когнітивних процесів, притаманних людині при формуванні та прийнятті класифікаційних рішень, тобто є максимально наближеними до механізму природного інтелекту;
- дозволяє оптимізувати в інформаційному розумінні просторово-часові параметри машинного навчання СППР, які впливають на її функціональну ефективність;
- методи ІЕІ-технології забезпечують високу функціональну ефективність СППР за умови перетину класів розпізнавання, що дає змогу

використовувати її в практичних задачах контролю, діагностування та керування у всіх галузях соціально-економічної сфери діяльності;

- оскільки вирішальні правила будуються за результатами машинного навчання в рамках геометричного підходу, то вони є практично інваріантні до проблеми багатовимірності словника ознак і алфавіту класів розпізнавання. .

Таким чином, аналіз тенденції розвитку методів інтелектуального аналізу даних вказує на те, що вже розроблено безліч алгоритмів, які здебільшого носять науково-методологічний характер, і характеризується специфікою застосування. Тому виникає гостра потреба вдосконалення та розроблення нових універсальних методів в рамках перспективного функціонального підходу до моделювання когнітивних процесів формування та прийняття класифікаційних рішень, найбільш наближених до природного інтелекту.

1.3 Формалізована постановка задачі інформаційного синтезу системи діагностування патологічних процесів, що навчається

Нехай відомий алфавіт класів розпізнавання $\{X_m^o \mid m=\overline{1,M}\}$, які характеризують M функціональних станів патологічного процесу. Для кожного класу за архівними даними клініко-лабораторних досліджень сформовано тривимірну вхідну навчальну матриця $\| y_{m,i}^{(j)} \|$ типу «об'єкт-властивість». Кожний рядок цієї навчальної матриці визначає структурований вектор ознак діагностування $\{y_{m,i}^{(j)} \mid i=\overline{1,N}\}$ класу розпізнавання X_m^o, де N – кількість ознак діагностування, а кожний стовпчик – випадкова навчальна вибірка ознаки діагностування $\{y_{m,i}^{(j)} \mid j=\overline{1,n}\}$, де n – обсяг вибірки. Таке представлення даних має свої

переваги: будь-яка сучасна мова програмування (C++, Java тощо) здатна оброблювати такі структури даних як тривимірні масиви, що значно полегшує роботу з ними, оскільки для них розроблена значна кількість алгоритмів та додаткових бібліотек. При цьому

$$N = N_1 + N_2,$$

де N_1 – кількість дійсних ознак розпізнавання, одержаних за результатами

клініко-лабораторних та імуногенетичних досліджень;

N_2 – кількість бінарних ознак розпізнавання, отриманих за результатами анамнезу.

Крім того, дано структурований вектор просторово-часових параметрів функціонування СППР, який у загальному випадку має структуру [1]

$$g =< g_1,...,g_{\xi_1},...,g_{\Xi_1},\ f_1,...,f_{\xi_2},...,f_{\Xi_2} >,\ \Xi_1 + \Xi_2 = \Xi, \qquad (1.1)$$

де $< g_1,...,g_{\xi_1},...,g_{\Xi_1} >$ – генотипні параметри функціонування, які впливають на параметри розподілу реалізацій образу;

$< f_1,...,f_{\xi_2},...,f_{\Xi_2} >$ – фенотипні параметри функціонування, які впливають на геометрію контейнерів класів розпізнавання, що відновлюються в радіальному базисі простору ознак.

При цьому відомі обмеження на відповідні параметри функціонування системи діагностування:

$$R_{\xi_1}(g_1,...,g_{\xi_1},...,g_{\Xi_1}) \le 0;\ \ R_{\xi_2}(f_1,...,f_{\xi_2},...,f_{\Xi_2}) \le 0.$$

Необхідно:

1. Визначити оптимальні значення параметрів функціонування СППР (1.1) $\{g_\xi^* \mid \xi = \overline{1, \Xi_1 + \Xi_2}\}$, які забезпечують максимум усередненого за алфавітом класів розпізнавання інформаційного критерію

$$\overline{E}^* = \frac{1}{M} \sum_{m=1}^{M} \max_{G_E \cap \{k\}} E_m^{(k)}, \tag{1.2}$$

де $E_m^{(k)}$ – інформаційний критерій оптимізації параметрів машинного навчання системи діагностування розпізнавати реалізації класу X_m^o, значення якого обчислено на k-му кроці машинного навчання;

G_E – допустима область визначення функції інформаційного критерію оптимізації, яка далі буде називатися робочою областю;

$\{k\}$ – впорядкована множина кроків навчання (відновлення контейнерів класів розпізнавання в радіальному базисі дискретного простору ознак).

2. Для апріорно класифікованого нечіткого розбиття $\tilde{\mathfrak{R}}^{|M|}$ побудувати шляхом допустимих перетворень в субпарацептуальному бінарному просторі ознак розпізнавання Хеммінга оптимальне (тут і далі в роботі в інформаційному розумінні) чітке розбиття класів розпізнавання $\mathfrak{R}^{|M|}$, на основі якого сформувати безпомилкові за навчальною матрицею вирішальні правила.

3. Розробити інформаційно-екстремальні метод машинного навчання СППР з гіперсферичними вирішальними правилами, що дозволить підвищити достовірність розпізнавання функціональних станів патологічного процесу.

4. На етапі екзамену з метою перевірки функціональної ефективності машинного навчання прийняти рішення про належність реалізації образу, що розпізнається, до одного із класів алфавіту $\{X_m^o\}$.

Таким чином, задача інформаційного синтезу здатної навчатися СППР для діагностування інфекційних патологічних процесів зводиться до оптимізації в процесі інформаційно-екстремального машинного навчання параметрів функціонування (1.1) за інформаційним критерієм (1.2) і до прийняття в режимі екзамену класифікаційного рішення за побудованими на етапі машинного навчання вирішальними правилами.

2. ОПИС МЕТОДУ ДОСЛІДЖЕНЬ

2.1 Основні принципи та положення інформаційно-екстремальної інтелектуальної технології аналізу даних

Згідно з працею [19] основна ідея машинного навчання у рамках інформаційно-еекстремальної інтелектуальнох технології (ІЕІ-технології) аналізу даних полягає в трансформації апріорного у загальному випадку нечіткого розбиття простору ознак у чітке розбиття класів еквівалентності шляхом ітераційної оптимізації параметрів функціонування системи розпізнавання. При цьому здійснюється цілеспрямовано пошук глобального максимуму багатоекстремальної функції статистичного інформаційного критерію в робочій (допустимій) області її визначення і одночасного відновлення оптимальних роздільних гіперповерхонь, що будуються в радіальному базисі бінарного простору ознак розпізнавання. Відмінністю методів ІЕІ-технології є те, що трансформація вхідного нечіткого розподілу реалізацій образів в чітке здійснюється в процесі оптимізації системи контрольних допусків, що приводить до цілеспрямованої зміни значень ознак розпізнавання і дозволяє побудувати безпомилкові за багатовимірною навчальною матрицею вирішальні правила.

Методи інформаційно-екстремального машинного навчання реалізовано крім відомих принципів системного аналізу на таких специфічних принципах [19,20]:

- принцип максимізації інформації, який обґрунтовується екстремальністю сенсорного сприйняття образу, що експериментально доведено вченими-фізіологами. Цей принцип реалізується шляхом введення додаткових інформаційних обмежень, що збільшує різноманітність об’єктів;
- принцип дуальності, який полягає в побудові найпростіших вирішальних правил за умови їх цілеспрямованого уточнення в міру

накопичення апостеріорної інформації з метою наближення до безпомилкових за навчальною вибіркою;

- принцип несумісності Л. Заде, який стверджує, що складність системи і точність у першому наближенні обернено пропорційні;
- принцип апріорної недостатності обґрунтування гіпотез (принцип Бернуллі-Лапласа), який стверджує, що за умов апріорної невизначеності доцільно розглядати апріорні гіпотези рівноймовірними, тобто прийняття рішень здійснюється за найгірших у статистичному розумінні умов;
- принцип рандомізації вхідних даних, який дозволяє досліджувати детерміновано-статистичні характеристики процесу;
- принцип зовнішнього доповнення, який обґрунтовує необхідність використання навчальної або контрольної (екзаменаційної) вибірки для підвищення достовірності рішень, що приймаються, та перевірки відповідності технічних характерристик системи заданим;

Побудова в процесі оптимізації параметрів машинного навчання вирішальних правил в рамках ІЕІ-технології здійснюється згідно з принципом відкладених рішень Івахненка О. Г. за багатоциклічною ітераційною процедурою пошуку максимального граничного значення усередненого за алфавітом класів розпізнавання інформаційного критерію оптимізації у вигляді

$$g_{\xi}^{*} = \arg\max_{G_{\xi}}\{\max_{G_{\xi-1}}\{\ldots\{\max_{G_1 \cap G_E}\frac{1}{M}\sum_{m=1}^{M}E_m\}\ldots\}\}, \qquad (2.1)$$

де E_m – інформаційний критерій оптимізації параметрів машинного навчання

системи розпізнавати реалізації класу X_m^o;

G_{ξ} – допустима область значень ξ-ї ознаки розпізнавання;

G_E – допустима область визначення функції інформаційного критерію оптимізації параметрів машинного навчання.

При цьому на алгоритм інформаційно-екстремального машинного навчання (2.1) накладаються обмеження:

$$\left(\forall X_m^o \in \tilde{\Re}^{|M|}\right)\left[X_m^o \neq \varnothing\right], \tag{2.2}$$

де $\tilde{\Re}^{|M|}$ – розбиття простору ознак на класи розпізнавання з потужністю $Card\,\tilde{\Re} = M$;

$$\left(\exists X_k^o \in \tilde{\Re}^{|M|}\right)\left(\exists X_l^o \in \tilde{\Re}^{|M|}\right)\left[X_k^o \neq X_l^o \to X_k^o \bigcap X_l^o \neq \varnothing\right]; \tag{2.3}$$

$$\left(\forall X_k^o \in \tilde{\Re}^{|M|}\right)\left(\forall X_l^o \in \tilde{\Re}^{|M|}\right)\left[X_k^o \neq X_l^o \to KerX_k^o \bigcap KerX_l^o = \varnothing\right], \tag{2.4}$$

де $Ker\,X_k^o$ – ядро класу розпізнавання X_k^o; $Ker\,X_l^o$ – ядро класу розпізнавання X_l^o, найближчого сусіда для класу розпізнавання X_k^o;

$$\left(\forall X_k^o \in \tilde{\Re}^{|M|}\right)\left(\forall X_l^o \in \tilde{\Re}^{|M|}\right)[X_k^o \neq X_l^o \to (d_k^* < d(x_k \oplus x_l))\,\&$$

$$\&(d_l^* < d(x_k \oplus x_l))], \tag{2.5}$$

де d_k^* – оптимальний радіус контейнера класу розпізнавання X_k^o;

d_l^* – оптимальний радіус контейнера класу розпізнавання X_l^o;

$d(x_k \oplus x_l)$ – кодова відстань між вектором x_k, усередненим за ансамблем реалізацій класу розпізнавання X_k^o і відповідним вектором x_l класу розпізнавання X_l^o;

$$\bigcup_{X_m^o \in \tilde{\Re}} X_m^o \subseteq \Omega_Б ; k \neq l; k,l,m = \overline{1,M} \ , \qquad (2.6)$$

де $\Omega_Б$ – бінарний простір Хеммінга.

Глибина інформаційно-екстремального машинного навчання характеризується кількістю параметрів функціонування системи, які оптимізуються за інформаційним критерієм. Кількість параметрів оптимізації визначається розмірністю вектора параметрів машинного навчання (1.1). При цьому внутрішній цикл реалізує так званий базовий алгоритм інформаційно-екстремального машинного навчання, який оптимізує геометричні параметри контейнерів класів розпізнавання, які відновлюються в радіальному базисі простору ознак розпізнавання.

Таким чином, за умови нечіткої гіпотези компактності структурованих векторів ознак розпізнавання , які далі будемо називати реалізаціями, основна ідея машинного навчання в методах IEI-технології полягає в послідовній адаптації вхідного математичного опису системи розпізнавання шляхом цілеспрямованої трансформації апріорних габаритів розкиду реалізацій образів з метою максимального їх захоплення контейнерами відповідних класів, які в процесі машинного навчання відновлюються в радіальному базисі простору ознак розпізнавання. Оптимальні контейнери за IEI-технологією забезпечують максимальну різноманітність між сусідніми класами, міра якої дорівнює максимуму інформаційного критерію оптимізації параметрів машинного навчання в робочій області визначення його функції. За отриманими в процесі машинного навчання оптимальними геометричними параметрами контейнерів класів розпізнавання будуються вирішальні правила, які дозволяють на екзамені приймати високодостовірні класифікаційні рішення.

Цілеспрямованість оптимізації просторово-часових параметрів функціонування системи розпізнавання в методах ІЕІ-технології здійснюється шляхом визначення на кожному кроці машинного навчання тенденції зміни асимптотичних точнісних характеристик класифікаційних рішень, які приймаються в процесі машинного навчання.

2.2 Інформаційні критерії оптимізації параметрів машинного навчання

Центральним питанням інформаційного синтезу здатної навчатися системи розпізнавання є оцінка функціональної ефективності процесу машинного навчання, яка визначає максимальну достовірність рішень, що приймаються на екзамені. Як критерії оптимізації параметрів машинного навчання в методах ІЕІ-технології можуть використовуватися різні критерії, які задовольняють таким властивостям інформаційної міри:

- інформаційна міра є величина дійсна і знакододатна як функція від імовірності;
- кількість інформації для детермінованих змінних ($p_i = 1$ або $p_i = 0$) дорівнює нулю;
- інформаційна міра має екстремум при значенні ймовірності $p_i = \frac{1}{m}$, де m – кількість якісних ознак розпізнавання.

Серед інформаційних мір для оцінки функціональної ефективності СППР, що навчається, перевагу слід віддавати статистичним логарифмічним критеріям, які дозволяють працювати з навчальними вибірками відносно малих обсягів. Серед таких критеріїв найбільшого використання знайшли ентропійні міри Шеннона та інформаційна міра Кульбака [1, 19, 20].

Подамо нормований ентропійний критерій оптимізації ьпараметрів машинного навчкання системи розпізнавати реалізації класу X_m^o у вигляді [19]:

$$E_m^{(k)} = \frac{I_m^{(k)}}{I_{\max}^{(k)}} = \frac{H_m^{(k)} - H_m^{(k)}(\gamma)}{H_m^{(k)}}, \tag{2.7}$$

де $I_m^{(k)}$ – кількість умовної інформації, що обробляється на k-му кроці машинного навчання системи розпізнавати реалізації класу X_m^o;

$I_{\max}^{(k)}$ – максимальна кількість умовної інформації, одержаної на k-му кроці машинного навчання;

$$H_m^{(k)} = -\sum_{l=1}^{M} p(\gamma_{l,k}) \log_2 p(\gamma_{l,k}) \tag{2.8}$$

апріорна (безумовна) ентропія, що існує на k-му кроці навчання системи розпізнавати реалізації класу X_m^o;

$$H_m^{(k)}(\gamma) = -\sum_{l=1}^{M}\sum_{m=1}^{M} p(\gamma_{l,k}) p(\mu_{m,k}/\gamma_{l,k}) \log_2 p(\mu_{m,k}/\gamma_{l,k}) - \tag{2.9}$$

апостеріорна (умовна) ентропія, що характеризує залишкову невизначеність після k-го кроку навчання системи розпізнавати реалізації класу X_m^o;

$p(\gamma_{l,k})$ – безумовна ймовірність прийняття на k-му кроці навчання гіпотези $\gamma_{l,k}$;

$p(\mu_{m,k}/\gamma_{l,k})$ – апостеріорна ймовірність прийняття на k-му кроці навчання рішення $\mu_{m,k}$ за умови, що прийнята гіпотеза $\gamma_{l,k}$.

Для двохальтернативної системи оцінок (M = 2) і рівноймовірних гіпотез, що характеризує згідно з принципом Бернуллі-Лапласа найбільш важкий у статистичному сенсі випадок прийняття рішень, після відповідної підстановки ентропій (2.8) і (2.9) у вираз (2.7) та заміни за формулою Байєса відповідних апостеріорних ймовірностей на апріорні ентропійний критерій набирає вигляду

$$E_m^{(k)} = 1 + \frac{1}{2}\left(\frac{\alpha_m^{(k)}(d)}{\alpha_m^{(k)}(d) + D_{2,m}^{(k)}(d)} \log_2 \frac{\alpha_m^{(k)}(d)}{\alpha_m^{(k)}(d) + D_{2,m}^{(k)}(d)} + \right.$$

$$+ \frac{\beta_m^{(k)}(d)}{D_{1,m}^{(k)}(d) + \beta_m^{(k)}(d)} \log_2 \frac{\beta_m^{(k)}(d)}{D_{1,m}^{(k)}(d) + \beta_m^{(k)}(d)} +$$

$$+ \frac{D_{1,m}(d)}{D_{1,m}^{(k)}(d) + \beta_m^{(k)}(d)} \log_2 \frac{D_{1,m}(d)}{D_{1,m}^{(k)}(d) + \beta_m^{(k)}(d)} +$$

$$\left. + \frac{D_{2,m}^{(k)}(d)}{\alpha_m^{(k)}(d) + D_{2,m}^{(k)}(d)} \log_2 \frac{D_{2,m}^{(k)}(d)}{\alpha_m^{(k)}(d) + D_{2,m}^{(k)}(d)} \right), \qquad (2.10)$$

де $\alpha_m^{(k)}(d)$ – помилка першого роду прийняття рішення на k-му кроці машинного навчання;

$\beta_m^{(k)}(d)$ – помилка другого роду;

$D_{1,m}^{(k)}(d)$ – перша достовірність;

$D_{2,m}^{(k)}(d)$ – друга достовірність; d – дистанційна міра, яка визначає радіуси гіперсферичних контейнерів, побудованих в радіальному базисі п ростору Хеммінга.

Оскільки точнісні характеристики є функціями відстані роздільної гіперповерхні від геометричних центрів контейнерів відповідних класів розпізнавання, то критерій (2.10) слід розглядати як нелінійний і взаємно-неоднозначний функціонал від точнісних характеристик, що потребує

знаходження в процесі иашинного навчання робочої (допустимої) області його визначення.

Розглянемо процедуру обчислення в практичних задачах критерію (2.10). Оскільки інформаційний критерій є функціоналом від точнісних характеристик, то при репрезентативному обсязі навчальної вибірки необхідно користуватися їх оцінками:

$$D_{1,m}^{(k)}(d)=\frac{K_{1,m}^{(k)}}{n_{\min}};\ \alpha_m^{(k)}(d)=\frac{K_{2,m}^{(k)}}{n_{\min}};\ \beta_m^{(k)}(d)=\frac{K_{3,m}^{(k)}}{n_{\min}};\ \ D_{2,m}^{(k)}(d)=\frac{K_{4,m}^{(k)}}{n_{\min}}, \qquad (2.11)$$

де $K_{1,m}^{(k)}$ – кількість подій, які означають належність “своїх” реалізацій класу

розпізнавання X_m^o ;

$K_{2,m}^{(k)}$ – кількість подій, які означають неналежність “своїх” реалізацій класу розпізнавання X_m^o;

$K_{3,m}^{(k)}$ – кількість подій, які означають належність “чужих” реалізацій класу розпізнавання X_m^o;

$K_{4,m}^{(k)}$ – кількість подій, які означають неналежність “чужих” реалізацій класу розпізнавання X_m^o;

$n_{\min}$ – мінімальний обсяг репрезентативної навчальної вибірки, який визначається за методом, запропонованим в праці [19].

Після підстановки відповідних позначень (2.11) у вираз (2.10) одержимо робочу формулу для обчислення в рамках IEI-технології ентропійного критерію машинного навчання системи розпізнавати реалізації класу X_m^o:

$$E_m^{(k)}=1+\frac{1}{2}\left(\frac{K_{1,m}^{(k)}}{K_{1,m}^{(k)}+K_{3,m}^{(k)}}\log_2\frac{K_{1,m}^{(k)}}{K_{1,m}^{(k)}+K_{3,m}^{(k)}}+\frac{K_{2,m}^{(k)}}{K_{2,m}^{(k)}+K_{4,m}^{(k)}}\log_2\frac{K_{2,m}^{(k)}}{K_{2,m}^{(k)}+K_{4,m}^{(k)}}+\right.$$

$$+\frac{K_{3,m}^{(k)}}{K_{1,m}^{(k)}+K_{3,m}^{(k)}}\log_2\frac{K_{3,m}^{(k)}}{K_{1,m}^{(k)}+K_{3,m}^{(k)}}+\frac{K_{4,m}^{(k)}}{K_{2,m}^{(k)}+K_{4,m}^{(k)}}\log_2\frac{K_{4,m}^{(k)}}{K_{2,m}^{(k)}+K_{4,m}^{(k)}}\Bigg). \qquad (2.12)$$

У праці [19] запропоновано модифікацію диференціальної інформаційної міри Кульбака, яка подається як добуток відношення правдоподібності на міру відхилень відповідних розподілів імовірностей. Для двохальтернативних апріорно рівноймовірних рішень модифікований критерій Кульбака, який обчислюється на k-му кроці машинного навчання системи розпізнавати реалізації класу X_m^o має вигляд

$$E_m^{(k)}=\log_2\left(\frac{2-(\alpha_m^{(k)}(d)+\beta_m^{(k)}(d))}{\alpha_m^{(k)}(d)+\beta_m^{(k)}(d)}\right)*\left[1-(\alpha_m^{(k)}(d)+\beta_m^{(k)}(d))\right]. \qquad (2.13)$$

Нормований критерій (2.13) можна подати у вигляді

$$E_{K,m}^{(k)}=\frac{E_{K\,m}^{(k)}}{E_{K\max}^{(k)}}, \qquad (2.14)$$

де $E_{K\max}^{(k)}$ – значення інформаційного критерію (2.13) при

$D_{1,m}^{(k)}(d)=D_{2,m}^{(k)}(d)=1$

і $\alpha_m^{(k)}(d)=\beta_m^{(k)}(d)=0$.

Робоча модифікація критерію (2.10) після відповідної підстановки оцінок (2.11) набуває вигляду

$$E=\frac{1}{n}\log_2\left\{\frac{2n+10^{-r}-\left[K_2^{(k)}+K_3^{(k)}\right]}{\left[K_2^{(k)}+K_3^{(k)}\right]+10^{-r}}\right\}\left[n-(K_2^{(k)}+K_3^{(k)})\right], \qquad (2.15)$$

де 10^{-r} – достатньо мале число, яке вводиться для уникнення поділу

на нуль;

r – число, яке рекомендується на практиці вибирати з інтервалу

$$1 < r \leq 3.$$

Таким чином, інформаційні критерії (2.10) і (2.13) є функціоналами як від точнісних характеристик класифікаційних рішень, так і від дистанційних критеріїв, що дозволяє їх вважати загальними критеріями валідності машинного навчання, оскільки вони є узагальненням відомих статистичних і детермінованих (дистанційних) критеріїв близькості.

2.3 Базовий алгоритм інформаційно-екстремального машинного навчання системи діагностування інфекційної патології

У рамках IEI-технології машинне навчання розглядається як оптимізація (тут і далі в інформаційному розумінні) параметрів функціонування системи шляхом пошуку глобального максимуму багатоекстремальної функції інформаційного критерію оптимізації. При цьому ступінь глибини інформаційно-екстремального машинного навчання визначається кількістю параметрів навчання, що оптимізуються. Внутрішній цикл в процедурі (2.1) оптимізує геометричні параметри контейнерів класів розпізнавання, які відновлюються на кожному кроці навчання в радіальному базисі простору діагностичних ознак. У випадку нормального розподілу векторів реалізацій класів розпізнавання навколо їх ядер доцільно відновлювати в просторі діагностичних ознак гіперсферичні контейнери. Для таких контейнерів параметром оптимізації є їх радіуси, які змінюються на кожному кроці навчання. При цьому оптимальне значення радіусу контейнера класу розпізнавання дорівнює його екстремальному значенню в робочій області G_E. В методах IEI-технології внутрішній цикл процедури

(2.1) реалізує так званий базовий алгоритм інформаційно-екстремального навчання, який на кожному кроці навчання обчислює значення інформаційного критерію оптимізації, здійснює пошук його глобального максимуму в робочій області визначення функції критерію та визначає оптимальні значення параметрів функціонування СППР.

Категорійна модель у вигляді спрямованого графу відображення множин, які застосовуються у процесі машинного навчання СППР за базовим інформаційно-екстремальним алгоритмом, показано на рис . 2.1 [19].

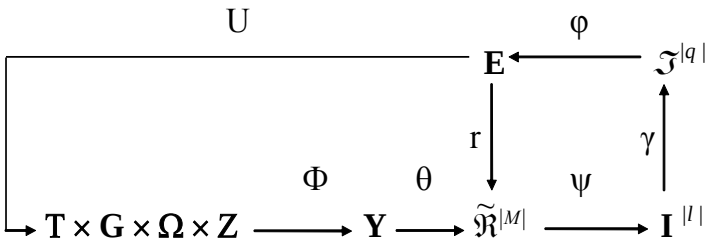


Рисунок 2.1 – Категорійна модель машинного навчання за базовим алгоритмом

На рис. 2.1 вхідний математичний опис задається у вигляді структури

$$I_B = <\boldsymbol{G}, \boldsymbol{T}, \Omega, \boldsymbol{Z}, \boldsymbol{Y}; \Phi>,$$

де T – множина моментів часу зняття інформації;

G – простір вхідних сигналів (факторів), які діють на СППР;

Ω – простір ознак розпізнавання;

Z – простір можливих станів, який визначається алфавітом класів розпізнавання;

Y – вибіркова множина, яка формує вхідну навчальну матрицю;

$\Phi: G \times T \times \Omega \times Z \to Y$ – оператор формування вибіркової множини Y, де декартовий добуток множин $G \times T \times \Omega \times Z$ розглядається як джерело інформації.

Оператор $\theta: Y \to \tilde{\Re}^{|M|}$ будує у загальному випадку нечітке розбиття $\tilde{\Re}^{|M|}$ простору ознак на класи розпізнавання, а оператор класифікації $\psi: \tilde{\Re}^{|M|} \to I^{|l|}$ перевіряє основну статистичну гіпотезу про належність реалізацій $\{x_m^{(j)} | j=\overline{1,n}\}$ нечіткому класу X_m^o і формує множину гіпотез $I^{|l|}$, де l – кількість статистичних гіпотез. Оператор $\gamma: I^{|l|} \to \Im^{|q|}$ шляхом оцінки статистичних гіпотез формує множину точнісних характеристик $\Im^{|q|}$, де $q=l^2$ – кількість точнісних характеристик. Оператор $\varphi: \Im^{|q|} \to E$ обчислює множину значень інформаційного критерію E, який є функціоналом точнісних характеристик. Контур оптимізації геометричних параметрів розбиття $\tilde{\Re}^{|M|}$ шляхом пошуку глобального максимуму інформаційного критерію оптимізації параметрів машинного навчання замикається оператором $r: E \to \tilde{\Re}^{|M|}$. Оператор U регламентує процес машинного навчання.

Вхідними даними базового алгоритму інформаційно-екстремального машинного навчання є апріорно класифікована нечітка навчальна матриця у вигляді тривимірного масиву реалізацій класів розпізнавання $\{y_m^{(j)} | m=\overline{1,M},\ j=\overline{1,n_m}\}$, система нижніх $\{A_{HK,i} | i=\overline{1,N_1}\}$ та верхніх $\{A_{BK,i} | i=\overline{1,N_1}\}$ контрольних допусків на діагностичні ознаки та рівень селекції ρ_m координат усередненого двійкового вектора ознак класу розпізнавання X_m^o, який по суті є рівнем квантування вхідних даних і за замовчуванням дорівнює $\rho_m = 0{,}5$.

Розглянемо основні етапи реалізації базового інформаційно-екстремального алгоритму машинного навчання за навчальною матрицею, яка складається як з кількісних діагностичних ознак, так і з категоріальних ознак, отриманих за результатами анамнезу пацієнтів.

1. Формування масиву бінарної навчальної матриці $\{x_{m,i}^{(j)} | m=\overline{1,M}, i=\overline{1,N}, j=\overline{1,n_m}\}$, елементи якої будемо визначати за правилом

$$x_{m,i}^{(j)} = \begin{cases} 1, \text{ if } \{A_{HK,i} < y_{m,i}^{(j)} < A_{BK,i}\} \wedge y_{m,i}^{(j)} \in \{S_k\}; \\ 0, \text{ if } \{A_{BK,i} \geq y_{m,i}^{(j)} \vee y_{m,i}^{(j)} \leq A_{HK,i}\} \wedge y_{m,i}^{(j)} \in \{S_k\}; \\ y_{m,i}^{(j)},\ y_{m,i}^{(j)} \notin \{S_k\}; \end{cases}$$

де $A_{HK,i}$ – нижній контрольний допуск для i-ї діагностичної ознаки усередненої реалізації y_1 базового класу X_1^o, який характеризує максимальну функціональну ефективність навчання СППР;

$A_{BK,i}$ – верхній контрольний допуск для i-ї діагностичної ознаки усередненого вектора y_1 базового класу X_1^o;

$y_{m,i}^{(j)}$ – значення i-ї ознаки в j-й реалізації класу X_m^o;

$\{S_k \mid k = \overline{1, N_1}\}$ – множина кількісних ознак розпізнавання.

2. Формування множини $\{x_m \mid m = \overline{1, M}\}$ усереднених векторів ознак для заданого алфавіту класів розпізнавання $\{X_m^o \mid m = \overline{1, M}\}$ за правилом [19]

$$\{x_{m,i} \mid m = \overline{1, M},\ i = \overline{1, N}\} = \begin{cases} 1, \text{ if } \dfrac{1}{n_m} \sum_{j=1}^{n} x_{m,i}^{(j)} > \rho_{m,i}; \\ 0, \text{ if else}, \end{cases}$$

де $x_{m,i}^{(j)}$ – значення i-ї ознаки в j-й реалізації бінарної навчальної матриці класу X_m^o;

$\rho_{m,i}$ – рівень квантування значень i-ї координати вхідних векторів-реалізацій класу X_m^o, який за замовчуванням дорівнює $\rho_{m,i} = 0{,}5$.

3. Виконується розбиття множини $\{x_m\}$ на пари найближчих (сусідніх) усереднених векторів-реалізацій

$$\{\Re_m^{|M|} = < x_m, x_c >\},$$

де x_c – усереднений вектор-реалізація сусіднього класу X_c^o, найближчого до класу X_m^o.

4. Обчислення кодових відстаней $d(x_m \oplus x_k^{(j)})$ від двійкового усередненого вектора x_m до реалізацій всіх класів $\{x_k^{(j)} \mid k = \overline{1,M},\ j = \overline{1,n_m}\}$.

5. Ініціалізація лічильника класів розпізнавання: $m := 0$.

6. $m := m + 1$.

7. Ініціалізація лічильника кроків зміни радіуса контейнера класу X_m^o: $d_m := 0$.

8. $d_m := d_m + 1$.

9. Обчислення точнісних характеристик $D_{1,m}$ та β_m за формулами (2.10). При цьому за сусідній клас X_c^o приймається сукупність найближчих до ядра класу X_m^o реалізацій

$$\{x_c^{(j)} \mid j = \overline{1,n_m}\} \subseteq \left\{ \bigcup_{k=1}^{M} X_k^o \setminus X_m^o \right\}.$$

10. Обчислення нормованого інформаційного критерію оптимізації навчання E_m за формулою (2.9).

11. Порівняння: якщо

$$d_m < d(x_m \oplus x_c),$$

то виконується перехід на крок 8, інакше – на крок 12.

12. Знаходження максимуму інформаційного критерію оптимізації в робочій області його визначення

$$E_m^* := \max_{G_{E_m} \cap G_{d_m}} E_m(d_m)$$

та оптимального радіусу контейнеру класу розпізнавання X_m^o за ітераційним алгоритмом максимізації інформаційного критерію оптимізації

$$d_m^* = \arg < \max_{G_{E_m} \cap G_{d_m}} E_m(d_m) >, \qquad (2.16)$$

де $E_m(d_m)$ – інформаційний критерій оптимізації параметрів навчання СППР, який є функцією від радіусу d_m контейнера класу X_m^o;

G_{E_m} – робоча область визначення функції критерію $E_m(d_m)$;

G_{d_m} – допустима область значень радіуса d_m контейнера класу X_m^o.

Саме процедуру (2.16) обчислення та пошуку глобального значення інформаційного критерію оптимізації в робочій області визначення його функції в методах IEI-технології прийнято називати базовим алгоритмом інформаційно-екстремального машинного навчання.

13. Порівняння: якщо $m < M$, то виконується перехід на крок 6, інакше – крок 14.

14. Запам’ятовуються оптимальні координати x_m^* і d_m^* вектора (1.1) параметрів функціонування СППР, за якими будуються вирішальні правила.

15. «ЗУПИН».

Таким чином, внутрішній контур процедури (2.1), який реалізує базовий алгоритм, має найбільшу обчислювальну трудомісткість , але він самостійно не забезпечує високу функціональну ефективність машинного навчання, оскільки вимагає оптимізації інших параметрів функціонування системи.

3. ІНФОРМАЦІЙНЕ, АЛГОРИТМІЧНЕ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ СИСТЕМИ ДІАГНОСТУВАННЯ ІНФЕКЦІЙНИХ ПАТОЛОГІЙ

3.1 Вхідний математичний опис системи діагностування стадій перебігу гострої кишкової інфекції

Алфавіт класів розпізнавання складався із трьох класів, які характеризували метод лікування гострої кишкової інфекції (ГКІ), викликаної умовно патогенними збудниками, а саме: контрольну групу осіб (здорові особи); групу осіб, для яких необхідне застосування базисної терапії з введенням внутрішньо пробіотика та клас розпізнавання, який характеризує одночасне призначення пробіотика та колоїдного срібла на тлі базисної терапії.

Вектори-реалізації класів розпізнавання представлялися у вигляді структурованої послідовності діагностичних ознак – результатів клініко-лабораторних досліджень мікробіоценозу кишечнику, рівня секреторного IgA, прозапального та протизапального цитокінів, гематологічних показників інтоксикації, що вводилися в систему діагностування.

Структурована реалізація функціонального стану патологічного процесу складалася із п’яти бінарних ознак – результатів анамнезу 120 пацієнтів та значень 19 кількісних діагностичних ознак, отриманих за результатами клініко-діагностичних досліджень цих пацієнтів:

1) Лейкоцитарний індекс інтоксикації;
2) Швидкість осідання еритроцитів ШОЕ (мм/год);
3) Лейкоцити – кількість лейкоцитів (10^9/л);
4) ГПІ – гематологічний показник інтоксикації;
5) ІЗЛК – індекс зсуву лейкоцитів;
6) Ілімф – лімфоцитарний індекс;

7) біфідобактерії (*lg* КУО / г);

8) Лактобацили (*lg* КУО / г);

9) Кишкова паличка зі слабко вираженими ферментними властивостями (*lg* КУО / г);

10) Загальна кількість кишкової палички (*lg* КУО / г);

11) Гемолізувальна кишкова паличка (*lg* КУО / г);

12) УПЕ – умовно патогенні ентеробактерії (*lg* КУО / г)

13) Золотистий стафілокок (*lg* КУО / г);

14) Гемолізувальний стрептокок (*lg* КУО / г);

15) Гриби роду Candida (*lg* КУО / г);

16) Секреторний імуноглобулін А (пг/л);

17) Інтерлейкін 1 бета (пг/л);

18) Інтерлейкін 4 (мг/л);

19) Патологічні мікроби сімейства «Кишкові».

Вхідна навчальна матриця формувалася за архівними даними, які були надані Сумською клінічною лікарнею інфекційних хвороб ім. Красовицького.

3.2 Інформаційно-екстремальне машинне навчання системи діагностування з оптимізацією контрольних допусків на діагностичні ознаки

Машинне навчання системи діагностування здійснювалося на прикладі визначення стадії патології ГКІ, викликаної умовно патогенними збудниками. Алфавіт класів розпізнавання складався із трьох класів. При цьому клас X_1^o характеризував контрольну групу осіб (здорові особи), клас X_2^o – групу пацієнтів, для яких необхідне комбіноване лікування з включенням до схеми колоїдного срібла і клас X_3^o – групу пацієнтів, для яких необхідне одночасне призначення пробіотика та колоїдного срібла на

тлі базисної терапії. Навчальні матриці класів мали по 40 реалізацій, кожна з яких складалася з 19 діагностичних ознак ($N_1 = 18,\ N_2 = 1$). При цьому вектори-реалізації класів подано у вигляді структурованої послідовності ознак розпізнавання, одержаних за результатами клінікр-лабораторних досліджень мікробіоценозу кишечнику, рівня секреторного IgA, протизапального цитокіну IL 1β, протизапального цитокіну IL 4 та інтегративних показників ендогенної інтоксикації.

Спочатку оптимізація геометричних параметрів контейнерів класів X_1^o, X_2^o і X_3^o при заданих контрольних допусках на ознаки розпізнавання здійснювалася за базовим алгоритмом, який полягав в реалізації вище наведеної в підрозділі 2.3 ітераційної процедури (2.16) пошуку глобального максимуму значення інформаційного критерію оптимізації Кульбака (2.15) в робочій області визначення його функції при відновленні в радіальному базисі гіперсферичних контейнерів класів розпізнавання із заданого алфавіту.

Привелені в підрозділі 3.3 результати машинного навчання характеризувалися невисокими максимальними значеннями інформаційного критерію (2.15) через неоптималне значення заданого параметра поля контрольних допускі на діагностичні ознаки, що вимагає збільшення глибини машинного навчання.

Розглянемо інформаційно-екстремальне машинне навчання системи діагностування ГКІ з оптимізацією система контрольних допусків на діагностичні ознаки. На рис. 3.2 показано двобічне симетричне поле допусків на значення i-ї ознаки $y_{m,i}$, $i = \overline{1, N_1}$ усередненої реалізації класу X_m^o.

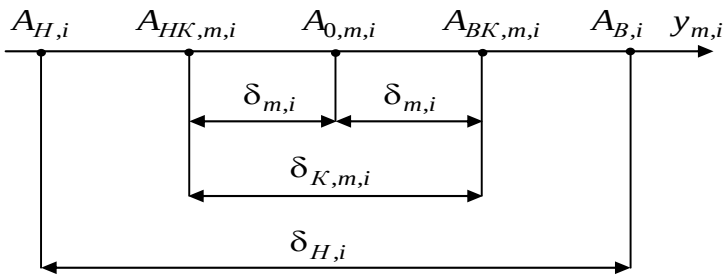


Рисунок 3.1 Двобічне симетричне поле допусків на значення ознаки розпізнавання

На рис. 3.1 прийнято такі позначення: $A_{0,m,i}$ – номінальне значення i-ї ознаки $y_{m,i}$; $A_{H,i}$, $A_{B,i}$ – нижній та верхній нормовані допуски відповідно; $A_{HK,m,i}$, $A_{BK,m,i}$ – нижній та верхній контрольні допуски відповідно; $\delta_{H,i}$ – нормоване поле допусків; $\delta_{K,m,i}$ – контрольне поле допусків; $\delta_{m,i}$ – параметр поля контрольних допусків.

При розв'язанні задачі інформаційного синтезу в рамках IEI-технології саме параметр $\delta = \delta_{m,i}, m = 1, i = \overline{1, N_1}$ розглядається як параметр оптимізації системи контрольних допусків на ознаки розпізнавання. При цьому знання на кожному кроці машинного навчання системи контрольних допусків на ознаки розпізнавання дозволяє здійснювати перехід із евклідового простору ознак в бінарний простір Хеммінга шляхом кодування значень аналогової ознаки розпізнавання на два рівні. Завдяки такому переходу з'являється можливість адаптувати сформовану робочу бінарну навчальну матрицю шляхом допустимих перетворень до побудови безпомилкових за її реалізаціями вирішальних правил.

У працях [1, 20] досліджувалася задача інформаційного синтезу здатної навчатися СППР із паралельно-послідовною оптимізацією контрольних допусків на ознаки розпізнавання. При цьому паралельна оптимізація полягає в одночасній зміні параметра δ поля контрольних допусків для всіх кількісних ознак розпізнавання, а послідовна – в почерговій зміні

контрольних допусків для кожної ознаки при заданих допусках для інших наступних ознак.

Відому структуру ітераційної процедури паралельної оптимізації системи контрольних допусків на ознаки розпізнавання подамо у вигляді [1]

$$\delta^* = \arg < \max_{G_\delta} \{ \frac{1}{M} \sum_{m=1}^{M} \left[\max_{G_{E_m} \cap G_{d_m}} E_m(d_m) \right] \} >, \qquad (3.4)$$

де $E_m(d_m)$ – інформаційний критерій оптимізації параметрів навчання СППР, який є функцією від радіусу d_m контейнера класу X_m^o;

G_δ – допустима область значень параметра $\delta = \delta_i, i = \overline{1, N_1}$ поля контрольних допусків;

G_{E_m} – робоча область визначення функції критерію $E_m(d_m)$;

G_{d_m} – допустима область значень радіуса d_m контейнера класу X_m^o.

Внутрішній цикл оптимізації системи контрольних допусків на ознаки розпізнавання реалізує базовий алгоритм навчання (3.3), основними функціями якого є обчислення значення інформаційного критерію оптимізації параметрів навчання СППР, пошук глобального максимуму в робочій (допустимій) області визначення його функції і оптимізація геометричних параметрів контейнерів класів розпізнавання. При цьому на кожному кроці зміни параметру функціонування СППР при поточних значеннях геометричних параметрів гіперсферичних контейнерів класів розпізнавання здійснюється обчислення усередненого за алфавітом класів розпізнавання інформаційного критерію оптимізації параметрів навчання СППР.

Алгоритм послідовної оптимізації системи контрольних допусків на ознаки розпізнавання, збіжність якого доведено в праці [19], полягає у наближенні глобального максимуму інформаційного критерію оптимізації до

граничного його значення в допустимій області значень функції критерію і має таку структуру:

$$\{\delta_i^* \mid i=\overline{1,N_1}\} = \arg < \overset{S}{\underset{s=1}{\otimes}} \left[\max_{G_{\delta_i}} \{ \frac{1}{M} \sum_{m=1}^{M} \left[\max_{G_{E_m} \cap G_{d_m}} E_m^{(s)}(d_m) \right] \} \right] >,$$

де G_{δ_i} – область допустимих значень параметра поля δ_i контрольних допусків для *i*-ї ознаки розпізнавання;

S – кількість прогонів ітераційної процедури послідовної оптимізації контрольних допусків на ознаки розпізнавання;

$\otimes$ – символ операції повторення.

Розглянемо кроки реалізації алгоритму (3.4) ітераційної паралельної оптимізації, при якій параметр δ поля контрольних допусків змінюється для всіх кількісних ознак розпізнавання одночасно.

Узагальнена схема алгоритму інформаційно-екстремального мащинного навчання системи розпізнавання визначається категорійною моделлю, показаною на рис. 3.2 [19].

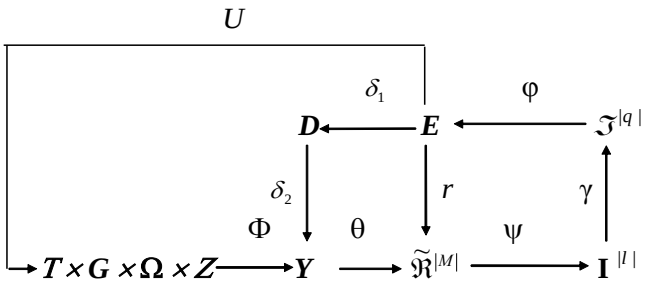


Рисунок 3.2 – Категорійна модель інформаційно-екстремального машинного навчання з оптимізацією контрольних допусків на ознаки розпізнавання

Відмінність категорійної моделі, показаної на рис. 3.2 від категорійної моделі машинногонавчання за базовим алгоритмом полягає у наявності

додаткового контура оптимізації контрольних допусків, який замикається ерез терм-множину D, елементами якої є впорядковані значення контрольних допусків на ознаки розпізнавання.

Вхідними даними алгоритму машинного навчання з паралельною оптимізацією контрольних допусків є: тривимірний масив навчальної матриці $\{y_{m,i}^{(j)} \mid m=\overline{1,M};\, i=\overline{1,N};\, j=\overline{1,n}\}$ реалізацій класів розпізнавання; нижні $\{A_{H,i} \mid i=\overline{1,N_1}\}$ і верхні $\{A_{B,i} \mid i=\overline{1,N_1}\}$ нормовані допуски на ознаки розпізнавання, які визначають область значень відповідних контрольних допусків.

Розглянемо схему інформаційно-екстремального машинного навчання системи діагностування ГКІ з оптимізацією система контрольних допусків на діагностичні ознаки.

1. Ініціалізація лічильника кроків зміни параметра поля контрольних допусків на ознаки розпізнавання $\delta := 0$.

2. $\delta := \delta + 1$.

3. Обчислення нижніх та верхніх контрольних допусків для всіх кількісних ознак розпізнавання за формулами

$$A_{HK,i} = y_{1,i} - \delta\frac{\delta_{H,i}}{100}; \tag{3.5}$$

$$A_{BK,i} = y_{1,i} + \delta\frac{\delta_{H,i}}{100}. \tag{3.6}$$

4. Реалізація базового алгоритму (2.16) машинного навчання за наведеною в підрозділі 2.3 схемою.

5. Порівняння: якщо $\delta \le \delta_{H,i}/2$, то виконується перехід на крок 2, інакше – перехід на крок 6.

6. Визначається максимальне усереднене за алфавітом класів розпізнавання значення $\overline{E}^*$ критерію (2.15) в робочій області визначення його функції та запам'ятовується екстремальне значення параметра δ^*.

7. Для оптимального параметра δ^* поля контрольних допусків за формулами (3.5) і (3.6) обчислюються оптимальні нижні $A^*_{HK,i}$ і верхні $A^*_{BK,i}$ контрольні допуски для всіх кількісних діагностичних ознак.

8. Запам'ятовуються оптимальні координати структурованого вектора (1.1) параметрів функціонування СППР:

$$g = <\{\delta^*_i = \delta^* \mid i = \overline{1, N_1}\}, x_m, d^*_m>.$$

9. «ЗУПИН».

За визначеними оптимальними геометричними параметрами контейнерів класів розпізнавання будуються вирішальні правила, які застосовуються при функціонуванні системи в режимі екзамену для перевірки функціональної ефективності машинного навчання і безпосередньо в робочому режимі для видачі діагностичного рішення, яке для лікаря-інфекціоніста має рекомендаційне значення. Для гіперсферичного класифікатора вирішальні правила в продукційній формі мають вигляд [20]

$$(\forall X^o_m \in \Re^{|M|})(\forall x^{(j)} \in \Re^{|M|})[if\ \ (\mu_m > 0)\ \&\ (\mu_m = \max\{\mu_m\})$$
$$then\ x^{(j)} \in X^o_m \quad else\ x^{(j)} \notin X^o_m],$$

(3.7)

де $x^{(j)}$ – вектор, що розпізнається;

μ_m – функція належності вектора $x^{(j)}$ класу розпізнавання X^o_m.

У виразі (3.7) функція належності для гіперсферичного контейнера класу розпізнавання X^o_m визначається за формулою [19]

$$\mu_m = 1 - \frac{d(x_m^* \oplus x^{(j)})}{d_m^*},$$

де x_m^*, d_m^* – оптимальні параметри машинного навчання: усереднена двійкова реалізація і радіус гіперсферичного контейнера класу розпізнавання X_m^o відповідно.

Таким чином, на екзамені визначається за вирішальними правилами (3.7) належність реалізації класу розпізнавання, що розпізнається, одному із класів із заданого алфавіту. При цьому вирішальні правила через малу обчислювальну трудомісткість відрізняються високою оперативністю і на відміну від нейроподібних структур практично інваріантні до вимірності простору ознак розпізнавання.

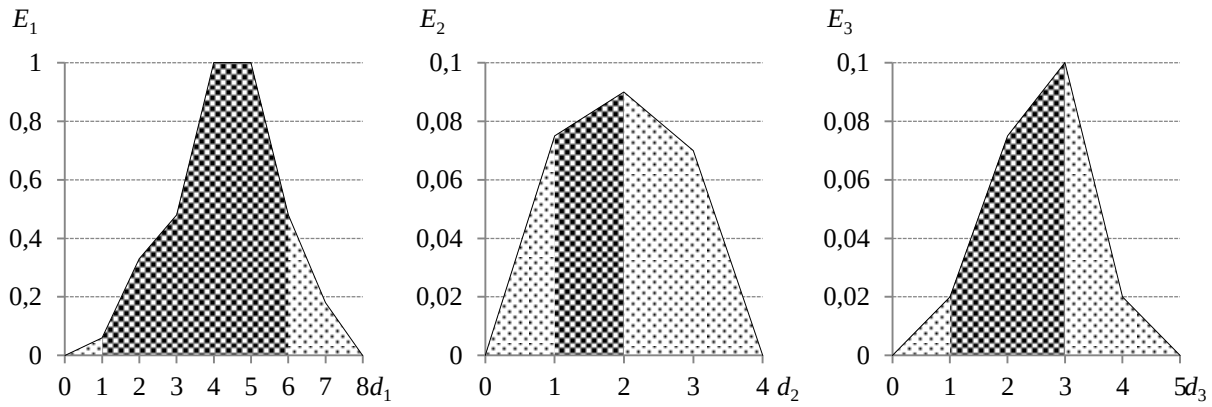
3.3. Результати фізичного моделювання

Програмна реалізація була створена на мові програмування C#, яка є об'єктно-орієнтованою мовою програмування з безпечною системою типізації для платформи .NET. Мова програмування C# близька до C++ і Java та має сувору статичну типізацію, підтримує поліморфізм, перевантаження операторів, вказівники на функції класу, атрибути, події, винятки та коментарі у форматі XML. Створена на основі своїх попередників C ++, Delphi, Module та Smalltalk мова C # виключаючи деякі моделі, які виявилися проблематичними в розробці програмного забезпечення. Так, C # не підтримує багаторазове успадкування класів в мові C ++ або тип виводу (відрізняється від Haskell). Основні методи програми машинного навчання наведено в таблиці 3.1

Таблиця 3.1 – Основні методи програми

Назва методу	Короткий опис
double[]searchAverage	Метод пошуку середнього значення
double[]searchLimit	Метод пошуку верхнього нижнього допуску
double[,]convertBinMatrix	Метод конвертації в бінарну матрицю
double[]SearchEtalVecBin	Метод пошуку еталонних векторів
double SearchMaxCount	Метод знаходження максимуму
double SearchMinCount	Метод знаходження мінімуму
double SearchCountXOR	Метод розрахунків оптимальних значень радіусів контейнерів класів розпізнавання
double[] SearchCountXORforEachLinesMatrix	Метод розрахуноків кодових відстаней між центрами класів та реалізаціями.
double[] SearchKFE	Метод пошуку значення КФЕ

На рис. 3.3 показано графіки залежності критерію (2.15) від радіусів гіперсферичних контейнерів класів розпізнавання, одержаних у процесі навчання СППР за вищенаведеним у підрозділі 2.3 базовим алгоритмом. Пр и цьому значення параметра поля контрольних допусків дорівнювало $\delta = 25$ у відсотках від номінального (усередненого) значення діагностичних ознак. На рис 3.3 (і далі в роботі) темною ділянкою позначено робочу (допустиму) область визначення функції інформаційного критерію оптимізації, в якій перша і друга достовірності діагностичних рішень, що приймаються, перевищують відповідно помилки першого і другого роду.



а *б* *в*

Рисунок 3.3 – Графіки залежності критерію (2.9) від радіусів гіперсферичних контейнерів класів розпізнавання: *а* – клас X_1^o; *б* – клас X_2^o; *в* – клас X_3^o

Аналіз рис. 3.3 показує, що максимальні значення інформаційного критерію оптимізації для класів X_1^o, X_2^o та X_3^o дорівнюють відповідно $E_1^* = 1$, $E_2^* = 0{,}09$ та $E_3^* = 0{,}1$, а оптимальні значення радіусів контейнерів відповідних класів розпізнавання – $d_1^* = 4$ (тут і надалі у роботі у кодових одиницях простору Хеммінга), $d_2^* = 2$ і $d_3^* = 3$. При цьому усереднене значення інформаційного критерію оптимізації виявилося невисоким і дорівнює $\overline{E}^* = 0{,}34$ через суттєвий перетин в просторі діагностичних ознак класів розпізнавання X_2^o і X_3^o.

З метою підвищення функціональної ефективності системи діагностування було реалізовано алгоритм машинного навчання з паралельною оптимізацією контрольних допусків, при якій на кожному кроці навчання всі допуски на діпгностичні ознаки змінювалися на задану величину одночасно. При цьому оптимальне значення параметра поля контрольних допусків дорівнювало $\delta^* = 60$ у відсотках від номінального (усередненого) значення діагностичних ознак.

На рис. 3.4 – 3.6 наведено графіки залежності інформаційного критерію оптимізації (2.15) від радіусів гіперсферичних контейнерів класів розпізнавання, одержаних за результатами паралельної оптимізації

контрольних допусків на діагностичні ознаки при оптимальній системіконтрольнихьдолпусків.

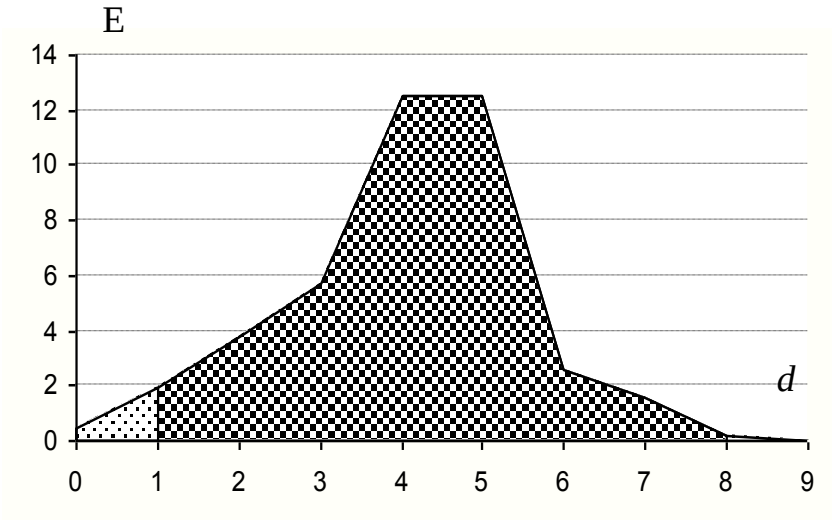


Рисунок 3.4 – Графік залежності інформаційного критерію від радіусів контейнера класу розпізнавання X_1^o

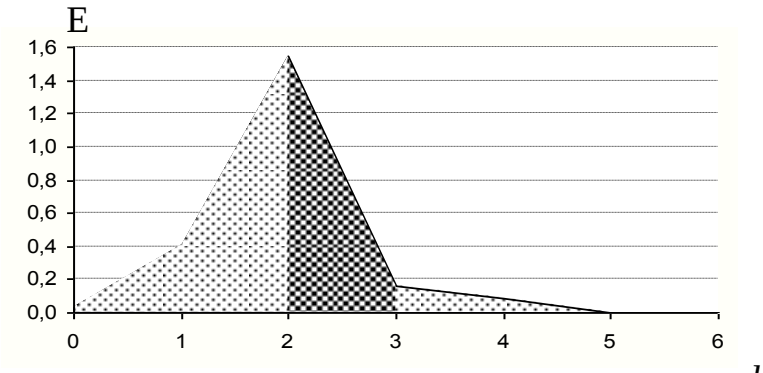


Рисунок 3.5 – Графік залежності інформаційного критерію від радіусів контейнера класу розпізнавання X_2^o

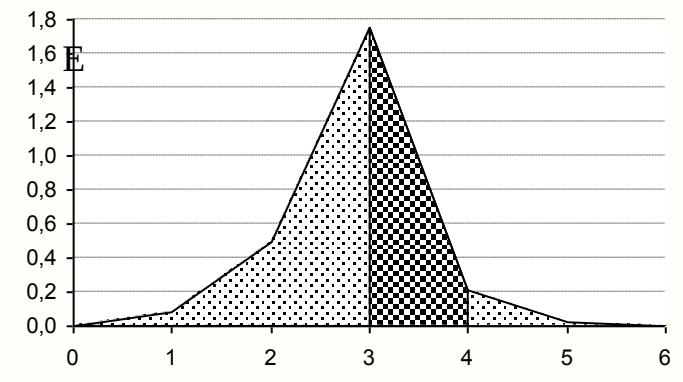


Рисунок 3.6 – Графік залежності інформаційного критерію від радіусів контейнера класу розпізнавання X_3^o

Аналіз рис. 3.5-3.6 показує, що оптимальні радіуси контейнерів дорівнюють $d_1^* = 4,\ d_2^* = 2$ і $d_3^* = 2$, а міжцентрові відстані між парами найближчих класів – $d(x_1 \oplus x_2) = 8,\ d(x_2 \oplus x_3) = 4$ і $d(x_3 \oplus x_2) = 4$ відповідно. При цьому середнє значення інформаційного критерію дорівнює $\overline{E}^* = 4{,}75$, що суттєво більше ніж на порядок перевершує його значення, отримане при реалізації базового алгоритму при неоптимальному параметрі δ поля контрольних допусків на діагностичні ознаки.

Таким чином, оптимізація системи контрольних допусків на ознаки розпізнавання дозволяє підвищити функціональну ефективність машинного навчання системи діагностування у порівнянні з базовим алгоритмом навчання, який було реалізовано при неоптимальній системі контрольних допусків.

Оскільки інформаційний критерій не досягає свого максимального граничного значення, яке обчислюється згідно з формулою (2.15) при $n = 30$ і $r = 2$, то розбиття простору ознак на класи розпізнавання все ще залишається нечітким, тобто існує перетин класів розпізнавання. Для підвищення функціональної ефективності машинного навчання необхідно збільшити його глибину шляхом оптимізації додаткових параметрів функціонування системи діагностування, включаючи параметри формування вхідного математичного опису. Крім того, при розширенні алфавіту класів розпізнавання доцільним є перехід на інформаційно-екстремальне машинне навчання системи діагностування за ієрархічною структурою даних.

ВИСНОВКИ

1. Розроблене алгоритмічне, інформаційне та програмне забезпечення машинного навчання системи діагностування інфекційної патології з оптимізацією системи контрольних допусків на діагностичні ознаки дозволяє підвищити функціональну ефективність машинного навчання системи діагностування у порівнянні з базовим алгоритмом навчання, який було реалізовано при неоптимальній системі контрольних допусків.

2. Оскільки інформаційний критерій оптимізації параметрів машинного навчання системи діагностування не досягає свого максимального граничного значення, то розбиття простору ознак на класи розпізнавання все ще залишається нечітким, тобто існує перетин класів розпізнавання.

3. Для підвищення функціональної ефективності машинного навчання необхідно збільшити його глибину шляхом оптимізації додаткових параметрів функціонування системи діагностування, включаючи параметри формування вхідного математичного опису.

4. При розширенні алфавіту класів розпізнавання доцільним є перехід на інформаційно-екстремальне машинне навчання системи діагностування за ієрархічною структурою даних.

СПИСОК ЛІТЕРАТУРИ

1. Довбиш А.С. Информационно-экстремальный алгоритм обучения системы диагностирования инфекционных патологий / А.С. Довбыш, А.А. Стадник, М.С. Руденко // Кибернетика и вычислительная техника. – 2013. – вып. 172. – С. 29-39.

2. Литвин А. А. Системы поддержки принятия решений в хирургии / А. А. Литвин, В. А. Литвин // Новости хирургии. – 2014. – Т. 2, № 1. – С. 96-100.

3. Шакало І. М. Ергономічні аспекти проектування систем реєстрації та обробки біомедичної інформації / І. М. Шакало, К. О. Чалий // Медична інформатика та інженерія. – 2009. – № 1. – С. 38-47.

4. Олексієнко М. М. Проблеми та перспективи впровадження інформаційних технологій в медичну практику / М. М. Олексієнко // Управління розвитком складних систем. – 2012. – Вип. 12. – С. 133-136.

5. Міхалевська Г. І. Основні концепції побудови експертних систем / Г. І. Міхалевська, В. Ц. Міхалевський // Collection of Scientific Papers of Applied Math and Computer Technologies Faculty of Khmelnytskyy National University. – 2010. – № 1(3). – С. 1-4.

6. Artificial Intelligence Techniques for Monitoring Dangerous Infections / [E. Lamma [et al.] // IEEE Transactions on Information Technology in Biomedicine. – 2006. – V. 10, I. 1. – P. 143-155.

7. Довбиш А.С. Информационно-экстремальный кластер-анализ в самообучающихся GRID-центрах / А.С. Довбыш, Саад Джулгам // 22-я Межд. Крымская конференция «СВЧ техника и телекоммуникационные технологии». Материалы конференции 10-14 сентября 2012 г. в 2 т.. – Севастополь, Украина: Вебер. – 2012. – Т.1. – С. 413-414.

8. Kldiashvil E. Grid Technologies for E-Health : Applications for Telemedicine Services and Delivery / E. Kldiashvil. – New York : Medical Information Science Reference, 2011. – 280 p.

9. Bellazzi R. Predictive Data Mining in Clinical Medicine : a Focus on Selected Methods and Applications / R. Bellazzi, F. Ferrazzi, L. Sacchi // Wiley Interdisciplinary Reviews : Data Mining and Knowledge Discovery. – 2011. – V. 1, I. 5. – P. 416-430.

10. Усков А. А. Экспресс-диагностика ОРВИ средствами нечетко-логической экспертной системы / А. А. Усков, М. В. Шипилов, В. В. Иванов // Программные продукты и системы. – 2011. – № 3. – С. 174-176.

11. Jiang J. Medical Image Analysis with Artificial Neural Networks / J. Jiang, P. Trundle, J. Ren // Computerized Medical Imaging and Graphics. – 2010. – V. 34, I. 8. – P. 617-631

12. Носенко Е. Н. Искусственные нейронные сети в медицинских исследованиях / Е. Н. Носенко, Д. Ю. Игнатов, Е. В. Зоркова // Перспективи медицини та біології. – 2011. – Т. III, № 1 (додаток). – С. 76-79.

13. Artificial Neural Networks – Methodological Advances and Biomedical Applications : [Ed. by K. Suzuki]. – InTech, 2011. – 362 p.

14. Зайченко Ю. П. Основи проектування інтелектуальних систем : навчальний посібник / Ю. П. Зайченко. – К. : Видавничий дім «Слово», 2004. – 352 с.

15. Анализ данных и процессов / [А. А. Барсегян [и др.]. – [3-е изд.]. – СПб. : БХВ-Петербург, 2009. – 512 с.

16. Graves D. Fuzzy C-Means, Gustafson-Kessel FCM, and Kernel-Based FCM: A Comparative Study / D. Graves, W. Pedrycz // Advances in Soft Computing. – 2007. – V. 41. – P. 140-149.

17. Shariati S. Comparison of ANFIS Neural Network with Several Other ANNs and Support Vector Machine for Diagnosing Hepatitis and Thyroid Diseases / S. Shariati, M. M. Haghighi // International Conference on Computer Information Systems and Industrial Management Applications, October 8-10, 2010 : Procedings. – Krackow, 2010. – P. 596-599.

18. Черезов Д. С. Обзор основных методов классификации и кластеризации данных / Д. С. Черезов, Н. А. Тюкачев // Вестник

Воронежского государственного университета. Серия : Системный анализ и информационные технологии. – 2009. – № 2. – С. 25-29.

19. Довбиш А.С. Інтелектуальні інформаційні технології в електронному навчанні / А.С. Довбиш, А.В. Васильєв, В.О. Любчак. – Суми: Видавництво СумДУ. – 2013. – 172 с.

20. Довбиш А.С. Основи проектування інтелектуальних систем: Навчальний посібник / А.С. Довбиш. – Суми: Видавництво СумДУ, 2009. – 171 с.

ДОДАТОК

```
public double[] searchLimit(double[] array, string which__one, double delta)
        {
            double[] retArray = new double[array.Length];
            if (which__one == "Up")
            {
                for (int i = 0; i < array.Length; i++)
                {
                    retArray[i] = array[i] + delta;
                }
            }
            if (which__one == "Down")
            {
                for (int i = 0; i < array.Length; i++)
                {
                    retArray[i] = array[i] - delta;
                }
            }
            return retArray;
        }
public double[] searchAverage(double[,] Matrix)
         {
            double z = 0;
            double[] meanMas = new double[Matrix.GetLength(0)];

            for (int i = 0; i < Matrix.GetLength(0); i++)
            {
                for (int j = 0; j < Matrix.GetLength(1); j++)
                {
                    z = z + Matrix[i, j];
                }
                meanMas[i] = z / 40d;
                z = 0;
            }
            return meanMas;
         }
```

```
public double[,] convertBinMatrix(double[,] matrix, double[] dopuskDownn, double[] dopuskUpp)
{
    double[,] Bin = new double[matrix.GetLength(0), matrix.GetLength(1)];

    for (int j = 0; j < matrix.GetLength(1); j++)
    {
        for (int i = 0; i < matrix.GetLength(0); i++)
        {

            if (matrix[i, j] < dopuskUpp[i] && matrix[i, j] > dopuskDownn[i])
            {
                Bin[i, j] = 1d;
            }
            else
            {
                Bin[i, j] = 0d;
            }
        }

    }
    return Bin;
}
```

```
public double[] SearchEtalVecBin(double[] array)
{
    double[] newMeanBin = new double[array.Length];
    for (int i = 0; i < array.Length; i++)
    {
        if (array[i] > 0.5d)
        {
            newMeanBin[i] = 1;
        }

        else
        {
            newMeanBin[i] = 0;
        }
    }
    return newMeanBin;
}
```

```
public double SearchMaxCount(double[] arr)
{
    double max = arr[0];
    for (int i = 1; i < arr.Length; i++)
    {
        if (arr[i] > max)
        {
            max = arr[i];
        }
    }
    return max;
}
```

```
public double[] SearchCountXORforEachLinesMatrix(double[] mainEtal, double[,] b
    {
        double[] array = new double[bin.GetLength(1)];
        int sum;
        for (int j = 0; j < bin.GetLength(1); j++)
        {
            sum = 0;
            for (int i = 0; i < bin.GetLength(0); i++)
            {
                if (mainEtal[i] != bin[i, j])
                {
                    sum++;
                }
            }
            array[j] = sum;
        }
        return array;
    }
```

```
public double[] SearchCountXORforEachLinesMatrix(double[] mainEtal, double[,] bin)
{
    double[] array = new double[bin.GetLength(1)];
    int sum;
    for (int j = 0; j < bin.GetLength(1); j++)
    {
        sum = 0;
        for (int i = 0; i < bin.GetLength(0); i++)
        {
            if (mainEtal[i] != bin[i, j])
            {
                sum++;
            }
        }
        array[j] = sum;
    }
    return array;
}
```