

Девіз «Пацієнт»

Науковий напрям  
«Інформатика, обчислювальна техніка та автоматизація»

**ІНТЕЛЕКТУАЛЬНА КОМП'ЮТЕРИЗОВАНА СИСТЕМА  
ДІАГНОСТУВАННЯ ІНФЕКЦІЙНОГО ЗАХВОРЮВАННЯ**

2012

## ЗМІСТ

ВСТУП.....	3
1 АНАЛІЗ ПРОБЛЕМИ ТА ПОСТАНОВКА ЗАДАЧІ .....	4
1.1 Короткий опис методів діагностування.....	4
1.2 Аналіз методів машинного навчання в теорії розпізнавання образів .....	5
1.3 Постановка задачі та завдання роботи .....	8
2 АНАЛІЗ ТА СИНТЕЗ СППР, ЩО НАВЧАЄТЬСЯ.....	11
2.1 Математичні моделі СППР, що навчається .....	17
2.2. Алгоритм навчання СППР для керування процесом діагностування стану пацієнта.....	11
2.3 Алгоритм функціонування СППР у режимі екзамену .....	15
3 ІНФОРМАЦІЙНЕ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ СППР, ЩО НАВЧАЄТЬСЯ.....	17
3.1 Формування вхідного математичного опису .....	<b>Ошибка! Закладка не определена.</b>
3.2 Короткий опис програмного забезпечення СППР, що навчається .....	18
ВИСНОВКИ .....	26
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ ІНФОРМАЦІЇ .....	27

## ВСТУП

Відомо, що діагностування захворювань пацієнтів здійснюється фахівцем на основі знання типових ознак, власного досвіду та професійної підготовки. При цьому не можна виключати впливу на результати діагностування і психоемоціонального стану самого лікаря. Тому розробка і впровадження в практичну медицину інтелектуальних комп'ютеризованих систем діагностування (КСД) і прогнозування перебігу і наслідків захворювань є нагальною потребою. Таким чином, тема наукової роботи є актуальною, оскільки вона спрямована на підвищення функціональної ефективності КСД гострого кишкового захворювання на його ранній стадії.

Основним функціональним блоком інтелектуальної КСД є система підтримки прийняття рішень (СППР), яка за умов апріорної невизначеності надає системі властивість адаптивності на основі машинного навчання та розпізнавання образів [1-3].

У роботі в рамках прогресивної інформаційно-екстремальної інтелектуальної технології (ІЕІ-технології) [4-5] розглядаються питання формування вхідного математичного опису, категорійного моделювання, оцінки функціональної ефективності та оптимізації параметрів функціонування здатної навчатися СППР, яка є складовою частиною КСД гострого кишкового захворювання.

## 1 АНАЛІЗ ПРОБЛЕМИ ТА ПОСТАНОВКА ЗАДАЧІ

### 1.1 Короткий опис методів діагностування

Гострі кишкові інфекції (ГКІ) є однією з актуальних медичних проблем. Протягом останнього десятиріччя в Україні склалася складна епідеміологічна ситуація, обумовлена сукупністю екологічних та соціально-економічних факторів, яка вплинула на загальне здоров'я населення. Щорічно наростає захворюваність на ГКІ в усіх вікових групах, яка за кількісними показниками стабільно посідає друге місце після гострих респіраторних вірусних інфекцій у структурі інфекційної патології. При цьому більшість з ГКІ викликані умовно патогенними мікроорганізмами (УПМ) родини *Enterobacteriaceae* (*Proteus mirabilis*, *Proteus vulgaris*, *Klebsiella pneumoniae*, *Enterobacter cloacae*, *Citrobacter freundii*, *Serratia marcescendes* та інше) [6].

Бактеріологічний метод є базовим для діагностики ГКІ. При цьому, не пізніше ніж за 2 години з моменту забору матеріалу від хворого, проводять посів випорожнень на спеціальні поживні середовища з наступною якісно-кількісною оцінкою отриманих результатів [7]. Але навіть при зміцненні матеріально-технічного забезпечення лабораторної служби, створенні бактеріологічних лабораторій у лікувально-профілактичних закладах у Сумській області протягом останніх трьох років, рівень етіологічної верифікації ГКІ сягає лише 58–59 % [8].

У динамічному спостереженні за хворим на ГКІ є необхідність вивчення змін мікробіоценозу кишечника так як глибина дисбіотичних змін прямо корелює з тяжкістю захворювання [9]. Це пов'язано з масивністю інвазії патогенів, десквамацією пошкодженого епітелію, рівнем збільшення популяційного рівня УПМ, наявністю супровідної патології ШКТ. При цьому патогенетичні механізми дії представників УПМ остаточно не з'ясовані, але звичайно, в розвитку патологічного процесу беруть участь фактори патогенності збудників, які мають токсичний вплив на макроорганізм та ендотоксинемія.

Важливо зазначити те, що ендогенна мікробна інтоксикація – явище постійне, так як у результаті загибелі різних мікроорганізмів відбувається виділення їх структурних компонентів, резорбція яких є мікробною складовою в розвитку системної ендогенної інтоксикації [10]. На інтенсивність даного процесу впливають колонізаційна резистентність кишечника, рівень природної резистентності організму хазяїна, інші фактори (перш за все, антимікробна та імуномодулююча терапія) [11]. При цьому раннє виявлення хворих із підвищеним ризиком розвитку ускладнень, можливим затяжним, рецидивуючим або хронічним перебігом інфекції, дозволяє лікарю своєчасно провести корекцію лікування.

Досягнуті успіхи в діагностуванні ГКІ пов'язані зі збільшенням обсягу інформації, яку обробляє особа, що приймає остаточне рішення – лікар-інфекціоніст. Але аналіз існуючих методів діагностики, що застосовуються у рутинній практиці, свідчить про відсутність автоматизованих комплексів для комплексної оцінки отриманих результатів. Тому нагальною потребою є розробка і впровадження в практичну охорону здоров'я систем підтримки прийняття рішень (СППР), що основані на машинному навчанні та розпізнаванні образів.

## 1.2 Аналіз методів машинного навчання в теорії розпізнавання образів та вибір методу досліджень

Становлення сучасної теорії аналізу і синтезу інтелектуальних СППР відбувається, головним чином, шляхом подальшого розвитку ідей і методів як розпізнавання образів, так і статистичної теорії прийняття рішень. Відомі детерміновані методи теорії розпізнавання образів характеризуються високою оперативністю, але низькою достовірністю розпізнавання [12]. Статистичні методи потребують наявності навчальних вибірок великих обсягів, що як правило ускладнено на практиці [13]. Також область використання статистичних методів при розпізнаванні образів обмежена тим, що її методи припу-

скають відомими умовні щільності ймовірностей, що для практичних задач отримати дуже важко. Структурні (лінгвістичні) методи розпізнавання характеризуються низькою оперативністю в режимі екзамену, тобто безпосереднього розпізнавання в робочому режимі, через необхідність формування образу, що розпізнається [14].

Основним недоліком багатьох методів які роблять неефективними їх застосування на практиці, є ігнорування перетину класів розпізнавання у просторі ознак розпізнавання, що має місце у практичних задачах.

Одними із перспективних методів аналізу і синтезу КСД, що навчаються, є методи, розроблені у рамках інформаційно-екстремальної інтелектуальної технології (ІЕІ-технологія), що ґрунтується на максимізації інформаційної спроможності системи шляхом введення додаткових інформаційних обмежень у процесі навчання [4,5].

Основні концептуальні положення ІЕІ-технології такі:

- технологія ґрунтується на прямій оцінці інформаційної спроможності СППР, що навчається;
- прийняття рішень здійснюється в рамках детерміновано-статистичного підходу шляхом побудови відносно простого детермінованого класифікатора, статистична корекція якого здійснюється в процесі навчання з метою підвищення достовірності рішень, що приймаються на екзамені;
- технологія базується на застосуванні гіпотези апріорної нечіткої компактності реалізацій образу, яка обумовлює перетин класів розпізнавання, що має місце в практичних задачах керування;
- технологія є об'єктно-структурованою, що дозволяє її розвивати для вирішення проблеми інформаційного синтезу широкого класу СППР;
- базується на вибіркового підході математичної статистики і орієнтована на застосування прийнятних з практичних міркувань обсягів репрезентативних навчальних вибірок;

– є універсальною для проектування здатної навчатися СППР будь-якої природи і дозволяє вирішувати як загальну, так і часткові задачі її інформаційного синтезу.

Розглянемо математичну модель навчання СППР за ІЕІ-технологією, яка включає як обов'язкову складову частину вхідний математичний опис у вигляді теоретико-множинної структури:

$$\Delta_A = \langle G, T, \Omega, Z, Y, \tilde{I}, \hat{O} \rangle,$$

де  $G$  – простір вхідних сигналів (факторів), які діють на СППР;  $T$  – множина моментів часу зняття інформації;  $\Omega$  – простір ознак розпізнавання;  $Z$  – простір можливих функціональних станів СППР;  $Y$  – навчальна матриця;  $\Pi: G \times T \times \Omega \rightarrow Z$  – оператор переходів, що відбиває механізм зміни станів СППР під дією внутрішніх і зовнішніх факторів;  $\Phi: G \times T \times \Omega \times Z \rightarrow Y$  – оператор формування вибіркової множини  $Y$  на вході СППР.

На рис. 1.1 показано базову у рамках ІЕІ-технології категорійну модель машинного навчання СППР у вигляді діаграми відображення множин, що використовуються в процесі навчання.

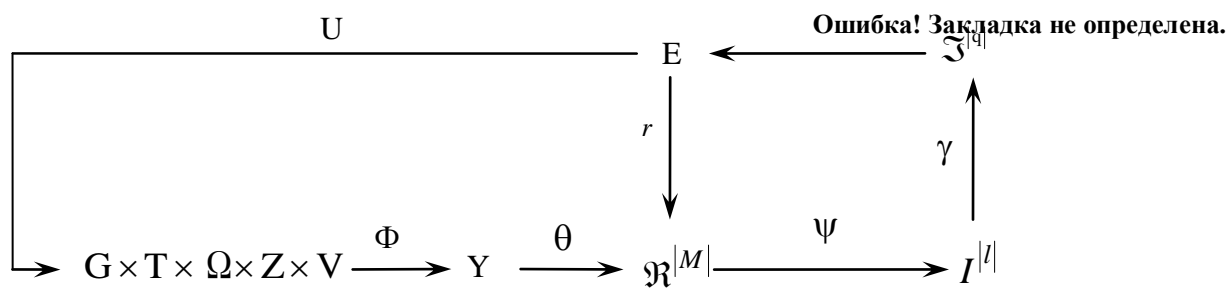


Рисунок 1.1– Структурна діаграма процесу навчання за ІЕІ-технологією

У діаграмі (рис. 1.1) оператор  $\theta: Y \rightarrow \tilde{\mathfrak{R}}^{|M|}$  будує у загальному випадку нечітке розбиття  $\tilde{\mathfrak{R}}^{|M|}$  простору ознак на  $M$  класів розпізнавання. Оператор

класифікації  $\psi: \tilde{\mathfrak{R}}^{M|} \rightarrow I^{l|}$  перевіряє основну статистичну гіпотезу про належність реалізацій  $\{x_m^{(j)} \mid j = \overline{1, n}\}$  класу  $X_m^o$  і формує множину рішень  $I^{l|}$ , де  $l$  – кількість статистичних гіпотез. Оператор  $\gamma$  шляхом оцінки одержаних рішень формує множину точнісних характеристик  $\mathfrak{S}^{q|}$ , де  $q = l^2$ . Оператор  $\varphi$  обчислює терм-множину  $E$  значень інформаційного критерію функціональної ефективності (КФЕ) навчання СППР, який є функціоналом від точнісних характеристик.

Таким чином, задача навчання у рамках ІЕІ-технології полягає у цілеспрямованій трансформації апріорно нечіткого розбиття простору ознак на класи розпізнавання у чітке розбиття шляхом пошуку глобального максимуму інформаційного КФЕ в робочій області визначення його функції.

### 1.3 Постановка задачі та завдання роботи

Розглянемо формалізовану постановку задачі інформаційного синтезу СППР у рамках ІЕІ-технології. Нехай сформовано алфавіт класів розпізнавання  $\{X_m^o \mid m = \overline{1, M}\}$ , що характеризують функціональний стан СППР, і відповідну навчальну багатовимірну (векторну) матрицю типу «об'єкт-властивість»  $\|y_{m,i}^j \mid i = \overline{1, N}, j = \overline{1, n}\|$ , де  $N, n$  – кількість ознак розпізнавання і векторів-реалізацій образу відповідно. При цьому рядок матриці  $\{y_{m,i}^j \mid i = \overline{1, N}\}$  утворює  $j$ -й вектор-реалізацію образу (далі може вживатися просто як реалізація), а стовпчик  $\{y_{m,i}^j \mid j = \overline{1, n}\}$  – навчальну вибірку з генеральної сукупності значень  $i$ -ої ознаки розпізнавання. Відомий структурований вектор параметрів функціонування СППР  $g = \langle x_m, d_m, \delta \rangle$ , де  $x_m$  – еталонний двійковий вектор (або реалізація) класу  $X_m^o$ ;  $d_m$  – радіус контейнера класу  $X_m^o$ , що відновлюється в радіальному базисі дискретного простору ознак розпізнавання,  $\delta$  – параметр поля контрольних допусків на



ознаки розпізнавання. При цьому задано такі обмеження: вершина вектора  $x_m$  визначає геометричний центр контейнера класу  $X_m^o$ ;  $d_m \in [0; d(x_m \oplus x_c) - 1]$ , де  $d(x_m \oplus x_c)$  – кодова відстань центра класу  $X_m^o$  від центру найближчого сусіднього класу  $X_c$  і  $\delta \in [0; \delta_H / 2]$ , де  $\delta_H$  – нормоване поле допусків, що визначає область значень параметра  $\delta$ .

Треба в процесі навчання СППР визначити оптимальні значення координат вектора  $g$ , що забезпечують максимальне значення усередненого за алфавітом  $\{X_m^o \mid m = \overline{1, M}\}$  КФЕ навчання СППР

$$\bar{E}^* = \frac{1}{M} \sum_{m=1}^M \max_{\{k\}} E_m, \quad (1.3.1)$$

де  $E_m$  – інформаційний КФЕ навчання СППР розпізнавати реалізації класу  $X_m^o$ ;  $\{k\}$  – впорядкована множина кроків навчання (відновлення контейнерів класів розпізнавання).

При функціонуванні СППР в режимі екзамену, тобто безпосереднього розпізнавання функціональних станів КСД, треба прийняти рішення про належність реалізації, що розпізнається, одному із класів сформованого на етапі навчання алфавіту  $\{X_m^o \mid m = \overline{1, M}\}$  і таким чином дефазифікувати функціональний стан хворого.

Основними завданнями розв'язання поставленої задачі є:

- сформулювати за експериментальними даними, отриманими в процесі обстеження хворих в Сумській обласній клінічній інфекційній лікарні ім. З.Й. Красовицького, вхідний математичний опис СППР у вигляді алфавіту класів розпізнавання  $\{X_m^o \mid m = \overline{1, M}\}$  і відповідної навчальної матриці  $\|y_{m,i}^j \mid i = \overline{1, N}, j = \overline{1, n}\|$ ;

- побудувати математичну модель функціонування СППР, що навчається з оптимізацією системи контрольних допусків (СКД), яка безпосередньо впливає на вхідний математичний опис;
- розробити алгоритми функціонування СППР в режимах навчання та екзамену;
- розробити програмне забезпечення СППР, що навчається;
- за результатами фізичного моделювання, тобто на реальних експериментальних даних, довести працездатність та надійність розробленого інформаційного та програмного забезпечення.

## 2 АНАЛІЗ ТА СИНТЕЗ СППР, ЩО НАВЧАЄТЬСЯ

2.1. Алгоритм навчання СППР для керування процесом діагностування стану пацієнта

Згідно з категорійною моделлю (рис.2.1) алгоритм навчання СППР у рамках ІЕІ-технології подамо як двоциклічну ітераційну процедуру пошуку глобального максимуму інформаційного КФЕ (1.3.1) в робочій (допустимій) області визначення його функції

$$\delta^* = \arg \max_{G_\delta} \{ \max_{G_E \cap G_d} \bar{E} \} >, \quad (2.1.1)$$

де  $G_\delta$  – допустима область значень параметра поля допусков  $\delta$ ;  $G_E$  – робоча область визначення функції критерію  $\bar{E}$ ;  $G_d$  – допустима область значень радіуса гіперсферичного контейнера класу розпізнавання.

Таким чином, як параметр навчання, що оптимізується у зовнішньому циклі, розглядається поле контрольних допусків  $\delta_{K,i}$  на ознаки розпізнавання, а як параметр навчання внутрішнього циклу оптимізації – радіус  $d_m$  гіперсферичного контейнера класу  $X_m^o$ , що відновлюється на кожному кроці навчання у радіальному дискретному просторі ознак.

Оптимізацію контрольних допусків на ознаки доцільно здійснювати за паралельно-послідовним алгоритмом, що забезпечує прийнятну оперативність та високу точність обчислення КФЕ. При цьому за алгоритмом паралельної оптимізації СКД на ознаки визначаються квазіоптимальні контрольні допуски, які для послідовного алгоритму приймаються як стартові.

Як вхідні дані використано три класи, що характеризують метод лікування гострої кишкової інфекції, викликані умовно-патогенними збудниками, а саме контрольну групу осіб (здорові особи), а також необхідність засто-

сування базисної терапії з введенням внутрішньо пробіотика, або базисної з введенням внутрішньо пробіотика та колоїдного розчину срібла; систему полів контрольних допусків  $\{\delta_{K,i} \mid i = \overline{1, N}\}$  на ознаки розпізнавання і рівень селекції (квантування)  $\{\rho_m\}$  координат еталонних векторів класів розпізнавання, який за замовчанням дорівнює  $\rho = 0,5$  для всіх класів розпізнавання.

Вектори-реалізації класів подано у вигляді структурованої послідовності ознак розпізнавання – клініко-лабораторних змін мікробіоценозу кишечника, рівня секреторного IgA, прозапальних та протизапальних інтерлейкінів, гематологічних показників, що вводяться лікарем.

Спочатку розглянемо узагальнені етапи реалізації інформаційно-екстремального алгоритму навчання СППР з паралельною оптимізацією контрольних допусків, коли відносний параметр поля допусків змінюється одночасно для всіх ознак розпізнавання:

- 1) обнуління лічильника кроків зміни параметра поля допусків  $\delta : l:=0$ ;
- 2)  $l:=l+1$ ;
- 3) на кожному кроці зміни параметра поля допусків обчислюються нижній  $A_{KH,i}[l]$  і верхній  $A_{KB,i}[l]$  контрольні допуски для всіх ознак розпізнавання за формулами

$$A_{KH,i}[l] = y_{1,i} - \delta[l]; \quad A_{KB,i}[l] = y_{1,i} + \delta[l], \quad (2.2.2)$$

де  $y_{1,i}$  – вибіркове середнє значення  $i$ -ї ознаки в навчальній матриці класу  $X_1^o$ .

- 4) формується бінарна навчальна матриця  $\|x_{m,i}^{(j)}\|$  за правилом

$$x_{m,i}^{(j)} = \begin{cases} 1, & \text{if } A_{KH,i} \leq y_{m,i}^{(j)} \leq A_{KB,i}; \\ 0, & \text{if } \textit{else}. \end{cases}$$

5) для класу  $X_m^o$  обчислюється двійковий еталонний вектор  $x_m$  за правилом

$$x_{m,i} = \begin{cases} 1, & \text{if } \frac{1}{n} \sum_{j=1}^n x_{m,i}^{(j)} > \rho_m; \\ 0, & \text{if } \text{else.} \end{cases}$$

де  $\rho_m$  – рівень селекції координат вектора  $x_m \in X_m^o$ , який за замовчуванням дорівнює  $\rho_m = 0,5$ .

6) для множини векторів  $\{x_m\}$  формується структурована множина елементів попарного розбиття  $\{\mathfrak{R}_m^{|2|} = \langle x_m, x_c \rangle\}$ , що задає план навчання. Тут  $x_c$  – еталонний вектор найближчого параметричного класу  $X_c^o$ .

7) обчислюється для кожного параметричного класу  $X_m^o$  значення інформаційного КФЕ навчання СППР розпізнавати його реалізації. Як КФЕ може розглядатися будь-яка статистична інформаційна міра [7,8]. Наприклад, для двохальтернативних рішень і рівноймовірних гіпотез застосуємо робочу модифікацію критерію Кульбака [5]

$$E = \frac{1}{n} \log_2 \left\{ \frac{2n + 10^{-r} - [K_2^{(k)} + K_3^{(k)}]}{[K_2^{(k)} + K_3^{(k)}] + 10^{-r}} \right\} [n - (K_2^{(k)} + K_3^{(k)})], \quad (2.2.3)$$

де  $K_{1,m}^{(k)}$  – кількість подій, які означають належність реалізацій образу контейнеру  $K_{1,m}^o$ , якщо дійсно  $\{x_1^{(j)}\} \in X_1^o$ ;  $K_{2,m}^{(k)}$  – кількість подій, які означають неналежність реалізацій контейнеру  $K_{1,m}^o$ , якщо дійсно  $\{x_1^{(j)}\} \in X_1^o$ ;  $K_{3,m}^{(k)}$  – кількість подій, які означають належність реалізацій контейнеру  $K_{1,m}^o$ , якщо вони насправді належать класу  $X_2^o$ ;  $K_{4,m}^{(k)}$  – кількість подій, які означають неналежність реалізацій контейнеру  $K_{1,m}^o$ , якщо вони насправді належать класу  $X_2^o$ ;  $n_{\min}$  – мінімальний обсяг репрезентативної навчальної вибірки,  $r$  – число цифр у мантисі значення критерію  $E_m^{(k)}$ .

8) обчислюється за формулами (1.3.1) і (2.1.3) усереднене значення критерію  $\bar{E}[l]$ .

9) якщо  $\delta \leq \delta_H / 2$ , то виконується пункт 2, інакше – пункт 10.

$$10) \bar{E}^*[l] := \max_{\{l\}} \bar{E}[l];$$

$$11) \delta^* := \arg \bar{E}^*[l];$$

12) для параметра  $\delta^*$  обчислюються за формулами (2.2.2) оптимальні нижні  $\{A_{HK,i}^*\}$  і верхні  $\{A_{BK,i}^*\}$  контрольні допуски на ознаки розпізнавання;

13) визначаються оптимальні радіуси контейнерів параметричних класів розпізнавання

$$d_m^* = \arg E_m^*[\delta^*],$$

де  $E_m^*[\delta^*]$  – максимальне значення КФЕ навчання СППР розпізнавати реалізації класу  $X_m^o$ , обчислене при оптимальному параметрі поля допусків  $\delta^*$ ;

14) ЗУПИН.

Одержані в процесі паралельної оптимізації параметри навчання використовуються як вхідні дані для алгоритму послідовної оптимізації контрольних допусків на ознаки розпізнавання. Алгоритм послідовної оптимізації поля контрольних допусків на ознаки розпізнавання здійснювався за такою ітераційною процедурою

$$\{\delta_{K,i}^*\} = \arg \left\{ \max_{G_{\delta_i}} \left\{ \max_{G_E} \left[ \max_{s=1}^S \max_{G_{d_1}} E_1^{(s)} \right] \right\} \right\}, i = \overline{1, N}, \quad (2.2.4)$$

де  $E_1^{(s)}$  – КФЕ навчання СППР розпізнавати реалізації базового класу  $X_1^o$ , обчислений за формулою (2.1.3) на  $s$ -му прогоні послідовної процедури оптимізації;  $G_{\delta_i}$ ,  $G_E$ ,  $G_{d_1}$  – області допустимих значень поля контрольних допусків для  $i$ -ї ознаки, критерію оптимізації і кодової відстані  $d_1$  відповідно;  $\otimes$  – символ операції повторення.

Таким чином, процес навчання СППР полягає в реалізації процедури пошуку глобального максимуму функції інформаційного критерію в робочій

області її визначення і ітераційного наближення цього максимуму до його граничного максимального значення, що забезпечує побудову безпомилкових за навчальною матрицею вирішальних правил, які геометрично характеризуються відновленими в процесі навчання оптимальними контейнерами параметричних класів розпізнавання.

## 2.2 Алгоритм функціонування СППР у режимі екзамену

Алгоритм екзамену за ІЕІ-технологією базується на аналізі значень функції належності, яка обчислюється для кожної реалізації, що розпізнається, і має, наприклад, до контейнеру класу  $X_m^o$  такий простий вигляд

$$\mu_m = 1 - \frac{d(x_m^* \oplus x^{(j)})}{d_m^*}, \quad (2.3.1)$$

де  $d(x_m^* \oplus x^{(j)})$  – кодова відстань між еталонним вектором  $x_m^*$  і реалізацією класу, що розпізнається;  $d_m^*$  – оптимальний радіус контейнера класу  $X_m^o$ .

Алгоритм екзамену має такі вхідні дані:  $M$  – кількість класів, які СППР навчена розпізнавати;  $\{x_m^* \mid m = \overline{1, M}\}$  – масив еталонних двійкових векторів-реалізацій образу, які визначають центри відповідних оптимальних контейнерів класів розпізнавання, побудованих на етапі навчання;  $\{d_m^*\}$ ,  $\{\delta_{k,i}^* \mid i = \overline{1, N}\}$  – масиви оптимальних радіусів контейнерів і контрольних допусків відповідно, визначенні на етапі навчання;  $\{x^{(j)}\}$  – масив двійкових векторів-реалізацій образу, що розпізнається.

Розглянемо етапи реалізації алгоритму екзамену:

- 1) формування лічильника класів розпізнавання:  $m := m + 1$ ;
- 2) формування лічильника числа реалізацій, що розпізнаються:  $j := j + 1$ ;

- 3) обчислення кодової відстані  $d(x_m^* \oplus x^{(j)})$ ;
- 4) обчислення функції належності;
- 5) порівняння: якщо  $j \leq n$ , то виконується крок 2, інакше – крок 6;
- 6) порівняння: якщо  $m \leq M$ , то виконується крок 1, інакше – крок 7;
- 7) визначення класу  $X_m^o$ , до якого належить реалізація образу, напри-

клад, за умови  $\bar{\mu}_m^* = \max_{\{m\}} \bar{\mu}_m$ , де  $\bar{\mu}_m = \frac{1}{n} \sum_{j=1}^n \mu_{m,j}$  – усереднене значення функ-

цій належності для реалізацій класу  $X_m^o$ , або видача повідомлення: «Клас не визначено», якщо  $\bar{\mu}_m^* \leq c$ . Тут  $c$  – порогове значення.

Таким чином, алгоритми екзамену у рамках ІЕІ-технології є детермінованими і відрізняються відносно малою обчислювальною трудомісткістю, що дозволяє їх реалізовувати у реальному темпі часу.



## 3 ІНФОРМАЦІЙНЕ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ СППР, ЩО НАВЧАЄТЬСЯ

### 3.1 Вхідний математичний опис КСД, що навчається

Як вхідні дані використано три класи, що характеризують метод лікування гострої кишкової інфекції, викликані умовно патогенними збудниками, а саме: контрольну групу осіб (здорові особи); необхідність застосування базисної терапії з введенням внутрішньо пробіотика та клас, що характеризує одночасне призначення пробіотика та колоїдного срібла на тлі базисної терапії; систему полів контрольних допусків  $\{\delta_{K,i} | i = \overline{1, N}\}$  на ознаки розпізнавання і рівень селекції (квантування)  $\{\rho_m\}$  координат еталонних векторів класів розпізнавання, який за замовчанням дорівнює  $\rho = 0,5$  для всіх класів розпізнавання.

Вектори-реалізації класів подано у вигляді структурованої послідовності ознак розпізнавання – клініко-лабораторних змін мікробіоценозу кишечника, рівня секреторного IgA, прозапального та протизапального цитокінів, гематологічних показників інтоксикації, що вводяться лікарем.

Структурована реалізація функціонального стану технологічного процесу складалася із значень таких 19 ознак розпізнавання:

- 1) лейкоцитарний індекс інтоксикації
- 2) швидкість осідання еритроцитів ШОЕ (мм/год)
- 3) лейкоцити – кількість лейкоцитів ( $10^9$ /л)
- 4) гематологічний показник інтоксикації ГПІ
- 5) індекс зсуву лейкоцитів ІЗЛК
- 6) лімфоцитарний індекс Ілімф
- 7) біфідобактерії ( $lg$  КУО / г)
- 8) лактобацили ( $lg$  КУО / г)
- 9) кишкова паличка зі слабо вираженими ферментними властивостями ( $lg$  КУО / г)

- 10) загальна кількість кишкової палички ( $lg$  КУО / г)
- 11) гемолізуювальна кишкова паличка ( $lg$  КУО / г)
- 12) умовно патогенні ентеробактерії УПЕ ( $lg$  КУО / г)
- 13) золотистий стафілокок ( $lg$  КУО / г)
- 14) гемолізуювальний 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»

### 3.2 Короткий опис програмного забезпечення СППР, що навчається

Для розроблення web-системи інформаційного забезпечення СППР, що навчається, в рамках ІЕІ-технології реалізовано наступні класи:

1.) Клас «*matrX*» – реалізовані функції зчитування та виведення на екран навчальних матриць та матриці екзамену; визначення математичного сподівання, мінімального та максимального значень ознак розпізнавання, значень нижнього та верхнього допусків на ознаки розпізнавання за паралельно-послідовним алгоритмом оптимізації контрольних допусків.

Функція *calcSKDParalel(\$h)* виконує алгоритм паралельної оптимізації контрольних допусків на ознаки розпізнавання:

```
function calcSKDParalel($h)
{
for ($i=0; $i<$_POST["m"]; $i++)
    if ($this->bOR[$i]==0) {
        $this->an[$i]=$this->ms[$i]-($this->ms[$i]/100)*$h;
        $this->av[$i]=$this->ms[$i]+($this->ms[$i]/100)*$h;
    }
}
```

Функція *calcSKDPosl(\$step)* виконує алгоритм послідовної оптимізації контрольних допусків на ознаки розпізнавання:

```
function calcSKDPosl($i, $h)
{
    $this->an[$i]=$this->ms[$i]-($this->ms[$i]/100)*$h;
    $this->av[$i]=$this->ms[$i]+($this->ms[$i]/100)*$h;
}
```

2.) Клас «*matrBinX*» – реалізовані функції формування та виведення на екран бінарних навчальних матриць; функції, що забезпечують функціонування ІС в режимі екзамену.

3.) Клас «*contejner*» – реалізовані функції формування та виведення на екран еталонних векторів; розрахунку матриці кодових відстаней між центрами класів та знаходження найближчого з них; розрахунку кодових відстаней між центрами класів та реалізаціями; обчислення точнісних характеристик та КФЕ за критеріями Кульбака та Шеннона.

Функція *Shenon\_INFk(\$k, \$t)* обчислює значення точнісних характеристик за критерієм Шеннона:

```
function Shenon_INFk($k, $t){
    $this->tochnKcharacteristics($k, $t);
    $this->fe[$t]=1+1/2*(calcLog($this->t_alpha[$t],$this->t_D2[$t])+calcLog($this->t_D1[$t],$this->t_beta[$t])+calcLog($this->t_beta[$t],$this->t_D1[$t])+calcLog($this->t_D2[$t],$this->t_alpha[$t]));
}
```

Функція *Kulbak\_INFk(\$k, \$t)* обчислює значення точнісних характеристик за критерієм Кульбака:

```
function Kulbak_INFk($k, $t){
    $this->tochnKcharacteristics($k, $t);
```

```

    $this->
    >fe[$t]=(1/$_POST["n"])*(log((2*$_POST["n"]+0.01-$this->
    >k2[$t]-$this->k3[$t])/($this->k2[$t]+$this->
    >k3[$t]+0.01))/log(2))*($_POST["n"]-$this->k2[$t]-
    $this->k3[$t]);
    }

```

У файлі «*BA.php*» реалізовано базовий алгоритм навчання ІС в рамках ІЕІ-технології.

```

<?
function calcBA($ro, $dX, &$dbX, &$cont){

    for ($k=0; $k<$_POST["k"]; $k++){
        $dbX[$k] = new matrBinX;
    }

    for ($k=0; $k<$_POST["k"]; $k++){
        $cont[$k] = new contejner;
    }

    for ($k=0; $k<$_POST["k"]; $k++)
        $dbX[$k]->getDatabX($dX[0], $dX[$k]);

    for ($k=0; $k<$_POST["k"]; $k++)
        $cont[$k]->getDataEV($dbX[$k], $ro);

    for ($k=0; $k<$_POST["k"]; $k++)
        $cont[$k]->makePara($k, $cont);

    for ($k=0; $k<$_POST["k"]; $k++)
        $cont[$k]->makeSk($k, $dbX);

    for ($k=0; $k<$_POST["k"]; $k++)
        $cont[$k]->calc_opt_rad($k);
}

?>

```

У файлі «*main.php*» реалізовано алгоритми паралельної та послідовної оптимізації СКД на ознаки розпізнавання, а також екзамени ІС в рамках ІЕІ-технології.

б) паралельна оптимізація СКД на ознаки розпізнавання:

```

for ($s=0; $s<=st; $s++){
    $dX[0]->calcSKDParalel($s);
    calcBA($_POST["ro"],$dX,$mdbX[$s],$mcont[$s]);
    getContDataIn-
fo($mcont[$s],$dc[$s],$mD1[$s],$mD2[$s],$ord[$s]);
}
calcMaxE($mdbX,$mcont,$dbX,$cont,$E1,$od);
for ($s=0; $s<=$st; $s++)
    for ($k=0; $k<$_POST["k"]; $k++){
        unset( $mdbX[$s][$k] );
        unset( $mcont[$s][$k] );
    }
if(isset($od)){
    $dX[0]->calcSKD($od);
    getContInfo($cont,$kfe1,$or1,$dc1,$dck1);
}

```

в) послідовна оптимізація СКД на ознаки розпізнавання:

```

for ($l=0; $l<3; $l++){
    for ($i=0; $i<$_POST["m"]; $i++){
        if($dX[0]->bOR[$i]==0){
            for ($s=0; $s<=st; $s++){
                $dX[0]->calcSKDPosl($i,$s);
                calcBA($_POST["ro"],$dX,$mdbX[$s],$mcont[$s]);
                getContDataInfo($mcont[$s],$dc[$s],
                    $mD1[$s],$mD2[$s],$ord[$s]);
            }
            calcMaxE($mdbX,$mcont,$dbX,$cont,$E2,$od2[$i]);
            for ($s=0; $s<=$st; $s++)
                for ($k=0; $k<$_POST["k"]; $k++){
                    unset( $mdbX[$s][$k] );
                    unset( $mcont[$s][$k] );
                }
            $dX[0]->calcSKDPosl($i,$od2[$i]);
        }
    }
    getContInfo($cont,$kfe2,$or2,$dc2,$dck2);
}

```

### 3.3. Результати фізичного моделювання

На базі вхідного математичного опису СППР, сформованого шляхом аналізу історії КСД стану пацієнта, було спочатку реалізовано за наведеною в підрозділі 2.1 схемою алгоритм навчання СППР з паралельною оптимізацією. На рис. 3.1 показано одержаний в процесі паралельної оптимізації СКД на ознаки розпізнавання графік залежності усередненого за формулами

(1.3.1) і (2.1.3) критерію Кульбака від параметра поля допусків  $\delta$ . При цьому при визначенні оптимальних симетричних контрольних допусків на кожній ітерації циклу збільшували значення параметру  $\delta$  на 1/100 від математичного очікування векторів базового класу  $X_1^o$ .

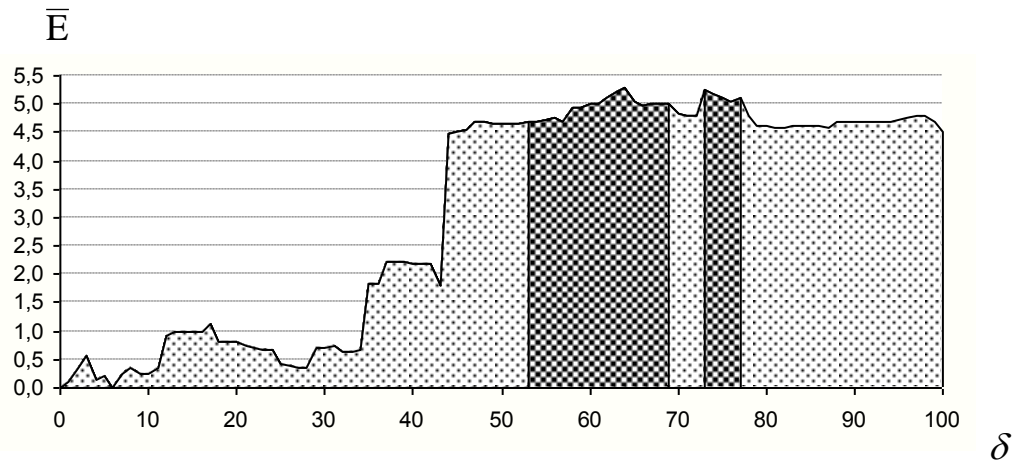
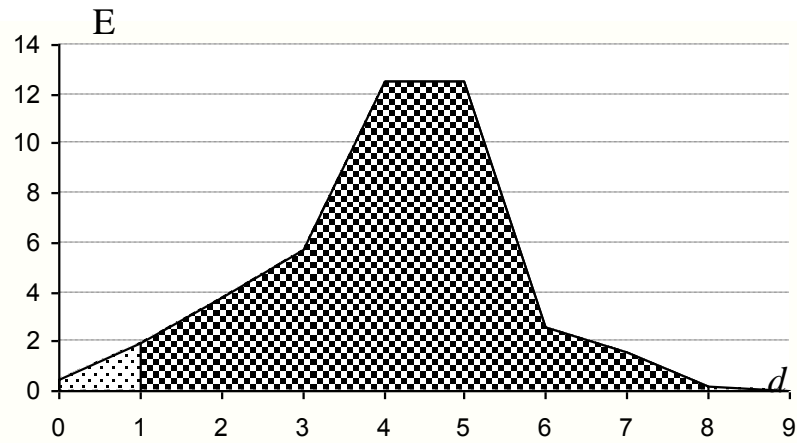
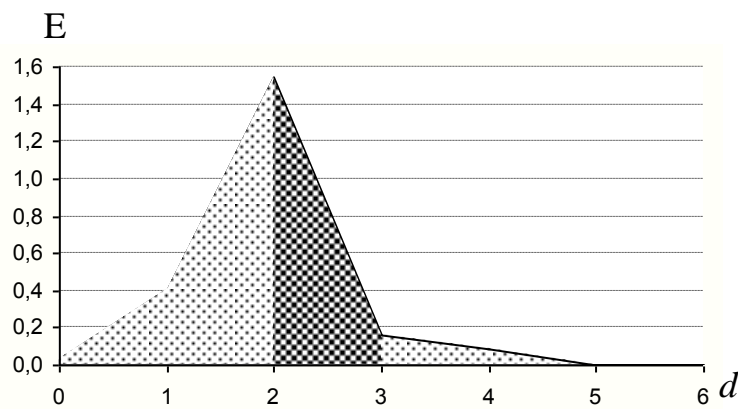
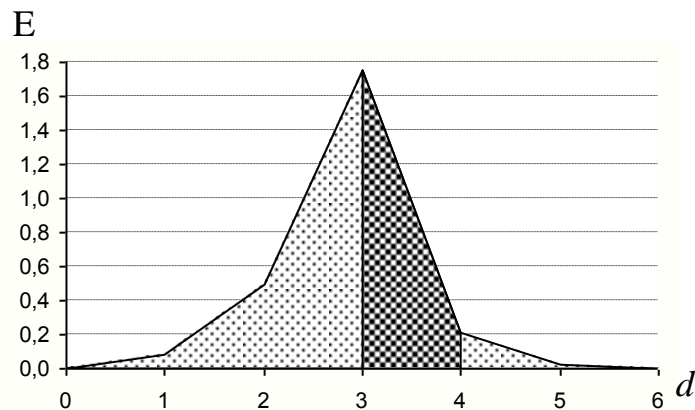


Рисунок 3.1 – Графік залежності усередненого КФЕ від параметра поля допусків  $\delta$  (delta)

На рис.3.1 темними ділянками позначено робочі області визначення функції (2.1.3), в яких виконуються нерівності  $D_{1,m} > 0,5$ ,  $D_{2,m} > 0,5$  і  $d_m^* < d(x_m \oplus x_c)$ . Остання нерівність вимагає виконання умови, щоб оптимальний радіус контейнера класу  $X_m^o$  був меншим за його міжцентрову відстань з найближчим класом. Аналіз рис. 3.1 показує, що квазіоптимальне значення параметра поля контрольних допусків дорівнює  $\delta^* = 64$  (у відносних одиницях поля допусків). При цьому максимальне усереднене за алфавітом класів розпізнавання значення КФЕ дорівнює  $\bar{E}^* = 5,283$ .

На рис. 3.2-3.4 показано графіки залежності КФЕ навчання СППР від радіусів контейнерів класів розпізнавання, що відновлюються в процесі навчання в радіальному базисі бінарного простору ознак.

Рисунок 3.2 – Залежність КФЕ від радіуса контейнера класу  $X_1^o$ Рисунок 3.3 – Залежність КФЕ від радіуса контейнера класу  $X_2^o$ Рисунок 3.4 – Залежність КФЕ від радіуса контейнера класу  $X_3^o$ 

Аналіз рис. 3.2-3.4 показує, що оптимальні радіуси контейнерів дорівнюють  $d_1^* = 4$ ,  $d_2^* = 2$  і  $d_3^* = 2$  (тут і далі у кодових одиницях), 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.

Оскільки значення КФЕ не досягає свого максимального граничного, яке обчислено згідно з формулою (2.1.3) при  $n=30$  і  $r=2$ , що свідчить про перетин класів розпізнавання, то для побудови безпомилкових вирішальних правил було реалізовано алгоритм послідовної оптимізації контрольних допусків на ознаки розпізнавання, в якому одержані при паралельній оптимізації квазіоптимальні контрольні допуски на ознаки розпізнавання розглядалися як стартові. Це дозволило суттєво підвищити оперативність алгоритму послідовної оптимізації контрольних допусків на ознаки розпізнавання, оскільки обчислення КФЕ навчання здійснювалося в робочій області визначення його функції. На рис. 3.5-3.7 показано графіки залежності ентропійного усередненого КФЕ від радіусів контейнерів відповідних класів розпізнавання.

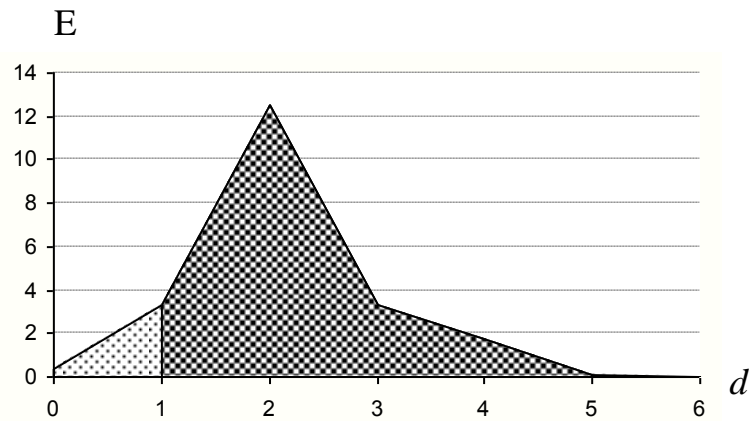


Рисунок 3.5 – Залежність КФЕ від радіуса контейнера класу  $X_1^o$  при паралельно-послідовній оптимізації СКД

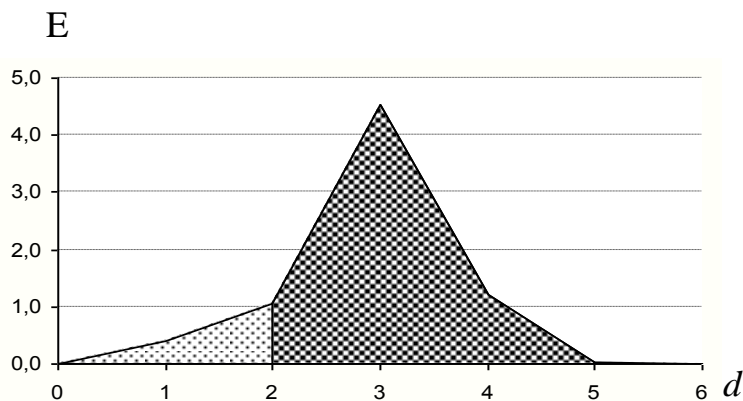


Рисунок 3.6 – Залежність КФЕ від радіуса контейнера класу  $X_2^o$  при паралельно-послідовній оптимізації СКД



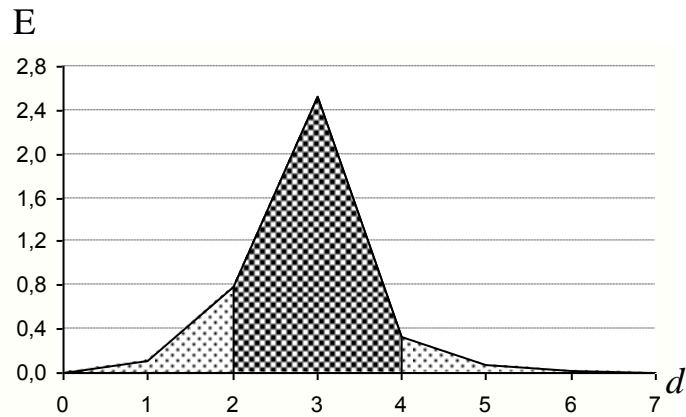


Рисунок 3.7 – Залежність КФЕ від радіуса контейнера класу  $X_3^o$  при паралельно-послідовній оптимізації СКД

Аналіз рис.3.5-3.7 показує, що в результаті паралельно-послідовній оптимізації СКД значення усередненого КФЕ за Кульбаком підвищилося у порівнянні з результатом, одержаним при паралельній оптимізації і дорівнює  $\bar{E}^* = 6,54$ . При цьому оптимальні радіуси контейнерів відповідно дорівнюють  $d_1^* = 2$ ,  $d_2^* = 3$  і  $d_3^* = 3$ , а міжцентрові відстані між парами найближчих класів –  $d(x_1 \oplus x_2) = 4$ ,  $d(x_2 \oplus x_1) = 4$  і  $d(x_3 \oplus x_2) = 5$  відповідно.

Подальші перспективи розвитку СППР з метою покращення інформаційної спроможності системи та підвищення достовірності розпізнавання під час екзамену полягають в застосуванні оптимізації рівня селекції координат двійкових еталонних векторів, словника ознак. При збільшенні потужності алфавіту класів розпізнавання з метою підвищення достовірності діагностування доцільним є перехід на ієрархічну структуру алгоритмів навчання і екзамену. Крім того, підвищення достовірності діагностування очікується при розширенні словника ознак розпізнавання за рахунок введення нових морфологічних ознак.

## ВИСНОВКИ

1. За результатами аналізу історій хвороб пацієнтів з ГКІ сформовано вхідний математичний опис здатної навчатися КСД у вигляді багатовекторної навчальної матриці для трьох класів розпізнавання, що характеризують стани пацієнтів

2. У рамках ІЕІ-технології розроблено інформаційне, алгоритмічне та програмне забезпечення здатної навчатися СППР для діагностування ГКІ, що дозволило надати КСД властивість адаптивності.

3. Досліджено вплив системи контрольних допусків на функціональну ефективність навчання КСД.

4. Одержані в роботі результати впроваджено в Сумській обласній інфекційній клінічній лікарні ім. З.Й. Красовицького і опубліковано у тезах доповіді Третьої Міжнародної науково-практичної конференції «Інтелектуальні системи в промисловості та освіті» (2-4 листопада 2011 р., м Суми).

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ ІНФОРМАЦІЇ

1. Ладанюк А.П. Автоматизація технологічних процесів і виробництв харчової промисловості: Підручник / А.П. Ладанюк, В.Г. Трегуб, І.В. Ельперін, В.Д. Цюцюра.– Київ: Аграрна освіта, 2001.–224 с.
2. Симанков В.С.. Адаптивное управление сложными системами на основе теории распознавания образов / В. С. Симанков, Е. В. Луценко.– Краснодар: Техн. ун-т Кубан. гос. технол. ун-та, 1999. — 318 с.
3. Кузьмин И.В. Основы теории информации и кодирования: Учебное пособие / И.В. Кузьмин, В.А. Кедрус. – К.: Вища школа, 1977.–415 с.
4. Краснопоясовський А. С. Інформаційний синтез інтелектуальних систем керування: Підхід, що ґрунтується на методі функціонально-статистичних випробувань / А.С. Краснопоясовський. Суми: Видавництво СумДУ, 2004. 261 с.
5. Довбиш А.С. Основи проектування інтелектуальних систем: Навчальний посібник /А.С. Довбиш.– Суми: Видавництво Сум ДУ, 2009.– 171 с.
6. Возианова Ж.И. Диареогенные кишечные палочки / Ж.И. Возианова // Сучасні інфекції. – 2008. – № 3. – С. 4–9.
7. Бактериологический посев (культурный метод) - <http://www.eurolab.ua/services/368/>.
8. Про епідемічну ситуацію в області у 2009 році - <http://www.ses.sumy.ua/index.php?id=309>.
9. Viswanathan V.K. Enteric infection meets intestinal function: how bacterial pathogens cause diarrhea / V.K. Viswanathan, K. Hodges, G. Hecht // Nat. Rev. Microbiol. – 2009. – № 7. – P. 110–119.
10. Фазылов В.Х. Системная эндотоксинемия и антиэндотоксиновая защита при острых кишечных инфекциях / В.Х. Фазылов, Д.Ф. Нестерова // Журнал инфектологии. – 2010. – Т. 2, № 4. – С. 120–121.
11. Гюлазян Н.М. Изменения профиля некоторых цитокинов при различных вариантах течения острых кишечных инфекций / Н.М. Гюлазян, О.Ф.

Белая, В.А. Малов [и др.] // Эпидемиология и инфекционные болезни. – 2008. – № 2. – С. 40–43.

12. Васильев В. И. Распознающие системы. Справочник. 2-е изд., перераб. и доп. / В.И. Васильев.– Киев: Наукова думка, 1983.– 422 с.

13. Vapnik V., *Statistical Learning Theory* / V. Vapnik. – John Wiley&Sons. New York, 1998. – 732 с.

14. Браверман Э. М. Структурные методы обработки эмпирических данных / Э.М. Браверман, И.Б. Мучник. – М.: Наука, 1983. – 464 с.