

**МІНІСТЕРСТВО ОХОРОНИ ЗДОРОВ'Я
НАЦІОНАЛЬНИЙ МЕДИЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ О. О. БОГОМОЛЬЦЯ**

**ФАРМАЦЕВТИЧНИЙ ФАКУЛЬТЕТ
КАФЕДРА АНАЛІТИЧНОЇ, ФІЗИЧНОЇ ТА КОЛОЇДНОЇ ХІМІЇ**

КВАЛІФІКАЦІЙНА ВИПУСКНА РОБОТА
на тему «**Прогнозування ступеня небезпечності/ризиків залишкових
розчинників у лікарських засобах методами хемометрії**»

Виконала: здобувачка вищої освіти 5 курсу
групи 8805

напряму підготовки 22 «Охорона здоров'я»
освітня програма «Фармація»

Калюженко Анастасія Володимирівна

Керівник: кандидат хімічних наук, доцентка
Пушкарьова Я.М.

Рецензент: проректор з науково-педагогічної
роботи, к.х.н., доцент, доцент кафедри хімічного
матеріалознавства хімічного факультету
Харківського національного університету імені
В.Н. Каразіна Пантелеймонов А. В.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ.....	3
ВСТУП	4
1. ЛІТЕРАТУРНИЙ ОГЛЯД	7
1.1. Залишкові розчинники в лікарських засобах.....	7
1.2. Методи аналізу та контроль наявності залишкових розчинників у лікарських засобах	8
1.3. Доцільність і ефективність методів хемометрії	10
1.4. Штучні нейронні мережі.....	11
<i>Висновки до розділу 1</i>	14
2. ЕКСПЕРИМЕНТАЛЬНА ЧАСТИНА	15
2.1. Масиви даних для дослідження	15
2.2. Розрахунок молекулярних дескрипторів.....	17
2.3. Прийоми зниження розмірності даних.....	23
2.3.1. Розрахунок критеріїв Краскела-Уолліса.....	23
2.3.2. Метод головних компонент.....	24
<i>Висновки до розділу 2</i>	28
3. РЕЗУЛЬТАТИ ТА ЇХ ОБГОВОРЕННЯ.....	29
3.1. Навчання ймовірнісної нейронної мережі	29
3.2. Прогнозування ступеня ризику/небезпечності залишкових розчинників у субстанціях, допоміжних речовинах і лікарських засобах.....	36
<i>Висновки до розділу 3</i>	40
ВИСНОВКИ.....	41
СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ	42
ДОДАТОК.....	47

**ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ,
СКОРОЧЕНЬ І ТЕРМІНІВ**

ГХ	газова хроматографія
ДФУ	Державна Фармакопея України
ШНМ	штучна нейронна мережа
ICH	Міжнародна конференція з гармонізації технічних вимог щодо реєстрації лікарських засобів для людини (The International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use)
ppm	part per million (1 мг/л, 1 мкг/мл)

ВСТУП

Актуальність теми. Органічні розчинники відіграють важливу роль у фармацевтичному виробництві, а саме при синтезі активних інгредієнтів і допоміжних речовин лікарських препаратів. Органічні розчинники можуть бути присутніми в добавках (консервантах, солубілізаторах, стабілізаторах) та в рецептурі лікарського засобу (в покритті таблетки або в матеріалі капсули, в розчинах настоянок, суспензій, емульсій, сиропів). Наявність органічних розчинників у невеликих кількостях може вплинути на ефективність, безпеку та стабільність фармацевтичної продукції.

Міжнародна конференція з гармонізації технічних вимог щодо реєстрації лікарських засобів для людини (ICH) ухвалила Настанову з регламентації залишкових розчинників, що встановлює межі вмісту розчинників, які можуть залишатися в діючих речовинах, допоміжних речовинах і лікарських засобах на різних стадіях виробництва та у готових лікарських препаратах.

На сьогодні є розчинники, що можуть становити інтерес для виробників діючих речовин, допоміжних речовин або лікарських засобів, але для яких немає достатніх токсикологічних даних, якими можна обґрунтувати допустиму добову дозу.

Таким чином, актуальність роботи обумовлено необхідністю створення теоретичного (комп'ютерного) підходу щодо прогнозування ступеня небезпечності/ризиків залишкових розчинників у лікарських засобах.

Мета дослідження: дослідити можливість застосування методів хемометрії для прогнозування ступеня небезпечності/ризиків залишкових розчинників у лікарських засобах за набором молекулярних дескрипторів.

Для досягнення поставленої мети необхідно було розв'язати такі **задачі:**

1) визначити набір молекулярних дескрипторів, що забезпечують правильну класифікацію залишкових розчинників відповідно до їх можливої загрози здоров'ю людини;

2) визначити оптимальну архітектуру ймовірнісної нейронної мережі, що забезпечує високу надійність класифікації залишкових розчинників за їх ступенем небезпечності/ризиків;

3) оцінити прогностичну силу запропонованої процедури прогнозування ступеня небезпечності/ризиків залишкових розчинників у лікарських засобах.

Предмет дослідження: параметри архітектури ймовірнісної нейронної мережі та закономірності її навчання при варіабельності вхідного набору даних.

Об'єкт дослідження: класифікація залишкових розчинників у субстанціях, допоміжних речовинах і лікарських засобах за ступенем ризику/небезпечності для здоров'я людини.

Методи дослідження: тест Краскела-Уолліса для встановлення найбільш інформативних молекулярних дескрипторів для надійної класифікації залишкових розчинників за їх ступенем ризику/небезпечності; метод головних компонент для зменшення розмірності отриманого масиву молекулярних дескрипторів; ймовірнісна нейронна мережа для встановлення класової приналежності залишкових розчинників за їх ступенем ризику/небезпечності.

Практичне значення отриманих результатів. Запропонована процедура прогнозування ступеня небезпечності/ризиків залишкових розчинників у лікарських засобах може використовуватися при плануванні та створенні нових синтетичних лікарських засобів.

За результатами роботи опубліковано статтю: Kaliuzhenko, A., & Pushkarova, Y. (2023). Application of artificial neural networks for solving pharmaceutical issues. *Grail of Science*, (24), 766-769.

Наукова новизна. Встановлено, що правильне навчання ймовірнісної нейронної мережі, а, відповідно, і правильну класифікацію залишкових розчинників за їх ступенем ризику/небезпечності, забезпечує набір із шести молекулярних дескрипторів: кількість акцепторів водневого зв'язку, логарифмічний коефіцієнт розчинності у воді, площа полярної поверхні, коефіцієнт форми, сума валентних ступенів, загальна валентна зв'язність.

Показано, що застосування ймовірнісної нейронної мережі забезпечує надійну класифікацію залишкових розчинників при широкому діапазоні значень відхилення функції активації.

1. ЛІТЕРАТУРНИЙ ОГЛЯД

1.1. Залишкові розчинники в лікарських засобах

Залишкові розчинники в фармацевтичних засобах це леткі органічні сполуки, які використовуються або утворюються в процесі виробництва фармацевтичних субстанцій чи готових лікарських засобів.

Хоча органічні сполуки є леткими сполуками, але після завершення технологічного процесу видаляються не повністю. У зв'язку з цим існує необхідність обов'язкового контролю їх вмісту в лікарських засобах.

Органічні розчинники використовують у фармацевтичному виробництві при синтезі активних інгредієнтів і допоміжних речовин лікарських препаратів. Органічні розчинники можуть бути присутніми в добавках (консервантах, солюбілізаторах, стабілізаторах) та в рецептурі лікарського засобу (в покритті таблетки або в матеріалі капсули, в розчинах настоянок, суспензій, емульсій, сиропів).

Особливістю органічних розчинників є їх токсичність незалежно від їх використання. Органічні розчинники небезпечні тим, що можуть акумулюватися в організмі.

Міжнародна конференція з гармонізації технічних вимог щодо реєстрації лікарських засобів для людини (ICH) ухвалила Настанову з регламентації залишкових розчинників, що встановлює межі вмісту розчинників, які можуть залишатися в діючих речовинах, допоміжних речовинах і лікарських засобах на різних стадіях виробництва та у готових лікарських препаратах. Встановлені кількісні межі органічних розчинників і обґрунтовано дані про безпеку та токсичність розчинників, запропоновані межі і норми виробничих процесів у фармацевтичному виробництві [1].

За рівнем ризику органічні розчинники поділяються на три класи [1, 2]:

- розчинники 1 класу – найбільш токсичні та/або екологічно небезпечні. Такі розчинники не повинні бути в ліках,
- розчинники 2 класу – середня категорія ризику. Такі розчинники

можуть бути в ліках в обмежених кількостях,

– розчинники 3 класу – найнижча категорія ризику. Такі розчинники мають обмеження в концентрації до 5000 ppm.

1.2. Методи аналізу та контроль наявності залишкових розчинників у лікарських засобах

Для визначення залишкових органічних розчинників відповідно до Державної Фармакопеї України (ДФУ) використовують газову хроматографію (ГХ) [2].

Газова хроматографія – метод розділення, в якому суміш речовин розділяють на окремі компоненти, використовуючи в якості рухомої фази інертний газ (аргон, гелій), нерухомою фазою слугує рідке або тверде покриття / сорбент в колонці; сполуки при розділенні можуть випаровуватися без розкладання, оскільки вони леткі та термостабільні.

До переваг газової хроматографії належать висока вибірковість та чутливість, легкість використання, проста пробопідготовка, легкість автоматизації, повна сумісність із сучасними нормативними вимогами.

Методи пробопідготовки при ГХ визначенні залишкових органічних розчинників [3, 4]:

1. Пряме введення розчину зразка – лікарську речовину або композицію розчиняють або екстрагують розчинником з високою температурою кипіння (диметилсульфоксид, диметилформамід, диметилацетамід, бензиловий спирт, етиленгліколь).

Переваги прямого введення зразка: простота, надійність, середня пропускна здатність.

Недоліки прямого введення зразка: низька чутливість, забруднення інжектора і колонки хроматографа, матричні ефекти, збільшення часу аналізу за рахунок необхідності чистки системи.

2. Парофазне концентрування: статичне парофазне концентрування та динамічне парофазне концентрування.

Регуляторні фармацевтичні органи пропонують статичне парофазне концентрування як найбільш прийнятний метод для визначення залишкових органічних розчинників в лікарських засобах.

Порівняння статичного парофазного концентрування та динамічного парофазного концентрування:

- статичне парофазне концентрування дозволяє визначати леткі сполуки,
- динамічне парофазне концентрування дозволяє визначати леткі та напівлеткі сполуки,
- динамічний метод є більш чутливим ніж статичний метод,
- у випадку динамічного парофазного концентрування сильно заважає матриця зразка,
- точність статичного методу вища за динамічний метод.

3. Твердофазна мікроекстракція – базується на сорбції аналіту на полімерне покриття фایбера (волокно) пристрою для твердофазної мікроекстракції.

Переваги парофазної твердофазної мікроекстракції:

- не заважає матриця зразка,
- висока чутливість визначення,
- метод дозволяє визначати леткі і напівлеткі сполуки,
- простий у використанні.

Метод парофазної твердофазної мікроекстракції витісняє метод динамічного парофазного концентрування при визначенні залишкових органічних розчинників у лікарських препаратах.

Етапи газохроматографічного процесу:

- випаровування зразка, його введення в інжектор і потрапляння в хроматографічну колонку,
- елюювання в колонці, що викликається потоком інертної газоподібної рухомої фази,
- транспортування сполуки через хроматографічну колонку рухомою

фазою (газом-носієм),

– розділення сполук та їх виявлення за допомогою відповідних детекторів.

При ГХ визначенні залишкових органічних розчинників найчастіше використовується полуменево-іонізаційний детектор – універсальний детектор на органічні сполуки. Перевагами полуменево-іонізаційного детектору є: низькі межі виявлення, широкий лінійний діапазон, простота експлуатації, загальна надійність і корисність, особливо при визначенні слідів органічних сполук [3, 4].

1.3. Доцільність і ефективність методів хемометрії

Хемометрія – це застосування статистичних і математичних методів до обробки аналітичних даних з метою збору та вилучення необхідної інформації для інтерпретації результатів аналізу.

Хемометрія є важливим та необхідним інструментом для роботи із багатовимірними (спектроскопічними, хроматографічними) масивами даних, а також для процесів калібрування, моделювання процесу, розпізнавання та класифікація образів, корекції сигналу та статистичного керування процесом [5, 6].

Є різні алгоритми та способи обробки і оцінки даних, і їх можна застосувати до вирішення ряду задач у медицині, фармації, контролі якості лікарських препаратів.

До найбільш популярних хемометричних методів належать [7]:

- метод головних компонент (principal component analysis, PCA),
- формальне незалежне моделювання аналогій класів (soft independent modelling of class analogy, SIMCA),
- дискримінантний аналіз (discriminant analysis),
- дерева класифікації та регресії (classification and regression trees, CART),
- метод опорних векторів (support vector machines, SVM),

- проєкція на латентні структури (projection on latent structures, PLS),
- штучні нейронні мережі (artificial neural networks, ANN).

Хемометричні методи застосовуються для аналізу даних виробничого процесу, випробувань контролю якості або вихідних даних інструментальних засобів з метою досягнення максимальної точності, точності та надійності. Хемометричні методи забезпечують швидкий кількісний аналіз фармацевтичних властивостей проміжних і готових лікарських форм, Хемометрія розвивалася паралельно з удосконаленням аналітичного обладнання та обчислювальних можливостей. Застосування хемометричних методів у фармацевтичній промисловості може варіюватися від встановлення специфікацій контролю якості сировини, порошків і лікарських форм до контролю різних процесів і етапів виробництва. Впровадження хемометричних методів з метою забезпечення загального контролю виробничого процесу передбачає використання аналітичних методів, здатних надавати точні результати простим і швидким способом [5, 6].

1.4. Штучні нейронні мережі

Штучні нейронні мережі (ШНМ) – це математичні моделі, сучасні обчислювальні системи, що складаються з елементарних одиниць обробки інформації (нейронів), певним чином з'єднаних один з другом і із зовнішнім середовищем.

Нейрони поєднуються в шари. Можна виділити три типи нейронів у залежно від виконуваних у мережі функцій [8]:

1) вхідні нейрони – служать лише для розподілу вхідних векторів і не виконують обчислень;

2) проміжні (приховані) нейрони – становлять основу нейронних мереж. Кожен вхідний вектор x_i множиться на відповідний ваговий коефіцієнт нейрона w_i та всі значення підсумовуються, визначаючи рівень активації нейрона NET:

$$NET = \sum_{i=1}^n x_i w_i + \theta, \quad (1.1)$$

де Θ – пороговий рівень нейрона (або зсув нейрона).

Далі сигнал NET перетворюється активаційною функцією F і формується вихідний нейронний сигнал OUT:

$$OUT = F(NET). \quad (1.2)$$

3) вихідні нейрони – представляють результати роботи нейронної мережі; перетворення в них також виконується за формулами (1.1) та (1.2).

Навчання ШНМ є двох типів: навчання з вчителем (контрольоване навчання), навчання без вчителя (неконтрольоване навчання). Процес навчання штучної нейронної мережі – це оптимізація архітектури нейронної мережі відповідно до вхідних даних з метою надійного та ефективного розв'язання поставленої задачі.

Сучасні алгоритми штучних нейронних мереж надзвичайно різноманітні за конфігураціями. Виділяють такі типи нейронних мереж [9, 10]:

- односпрямовані та двоспрямовані,
- рекурентні,
- мережі на основі радіальних базисних функцій та самоорганізовані карти.

Нейронні мережі класифікують за різними ознаками:

- типом вхідної інформації (аналогові та бінарні);
- типом навчання (нейронні мережі «з навчанням» та «без навчання»);
- характером налаштування вагових коефіцієнтів (нейронні мережі з статичними або динамічними зв'язками);
- топологією (одношарові та багатошарові мережі);
- характером зв'язків (мережі з прямими, перехресними, зворотними або латеральними зв'язками).

«Побудова» нейронної мережі складається з двох етапів:

1) вибір архітектури нейронної мережі – ініціалізація мережі, визначення числа нейронів, тип з'єднання нейронів, вид активаційних (передавальних) функцій;

2) навчання мережі – вибір способу навчання.

Застосування штучних нейронних мереж у фармацевтичній промисловості стає все більш активним та популярним. Наведемо кілька прикладів застосування алгоритмів штучних нейронних мереж для вирішення фармацевтичних завдань (Таблиця 1.1) [11–13].

Таблиця 1.1. Приклади застосування алгоритмів штучних нейронних мереж до вирішення фармацевтичних завдань [11]

Застосування	Приклади	Посилання
Прогнозування різних типів активності	Прогнозування протигрибкових властивостей	[14]
	Прогнозування антимікробної активності	[15]
	Прогнозування біологічної активності	[16]
Прогнозування токсичності сполук	Прогнозування токсичності фенолів	[17]
	Прогнозування токсичності сполук гідразину	[18]
	Прогнозування токсичності йонних рідин	[19]
Якісний аналіз (ідентифікація, кластеризація)	Класифікація фармацевтичних зразків з використанням «електронного язика»	[20]
	Класифікація олігопептидів і амінокислот за допомогою вольтамперометричного «електронного язика»	[21]

Висновки до розділу I

1. Контроль вмісту залишкових розчинників у діючих речовинах, допоміжних речовинах і лікарських засобах на різних стадіях виробництва та у готових лікарських препаратах є обов'язковим та є показником якості.

2. Зростає роль хемометричних методів, без яких неможлива обробка великих масивів багатовимірних даних. У зв'язку з цим, прагнучи високої надійності висновків якісного чи кількісного аналізу, дослідник повинен уважно вибирати не лише експериментальні методи, але й хемометричні засоби обробки результатів аналізу.

2. Штучні нейронні мережі завдяки адаптивній структурі та здатності до навчання успішно використовуються для вирішення різноманітних фармацевтичних задач.

2. ЕКСПЕРИМЕНТАЛЬНА ЧАСТИНА

2.1. Масиви даних для дослідження

У даній роботі для дослідження можливості прогнозування ступеня ризику/небезпечності залишкових розчинників у субстанціях, допоміжних речовинах і лікарських засобах були використані дані, що включені до Настанови з регламентації залишкових розчинників, ухваленою Міжнародною конференцією з гармонізації технічних вимог щодо реєстрації лікарських засобів для людини (ICH), та висвітлені у Державній Фармакопеї України [1, 2]. Масив даних містив:

– 59 розчинників, що за їхньою можливою загрозою здоров'ю людини класифіковані на три класи (таблиця 2.1): розчинники, використання яких слід уникати (клас 1); розчинники, використання яких слід обмежувати (клас 2); малотоксичні розчинники (клас 3);

– 9 розчинників¹, що можуть становити інтерес для виробників лікарських засобів чи допоміжних речовин, для яких відсутні дані з токсичності (таблиця 2.2).

Таблиця 2.1. Розчинники, для яких ступінь ризику/небезпечності відповідно до їх загрози здоров'ю людини відомі [2]

№ п\п	Назва розчинника	Клас розчинника
1	Анізол	3
2	Ацетон	3
3	Ацетонітрил	2
4	Бензол	1
5	1-Бутанол	3
6	2-Бутанол	3
7	Бутилацетат	3
8	<i>трет</i> -Бутилметиловий ефір	3
9	Гексан	2
10	Гептан	3
11	<i>N,N</i> -Диметилацетамід	2

¹ Петролейний ефір було виключено з переліку, оскільки це суміш насичених вуглеводнів, що унеможлиблює написання чіткої структурної формули та розрахунок молекулярних дескрипторів.

12	Диметилсульфоксид	3
13	<i>N,N</i> -Диметилформамід	2
14	1,2-Диметоксіетан	2
15	1,4-Діоксан	2
16	Дихлорметан	2
17	1,2-Дихлоретан	1
18	1,1-Дихлоретен	1
19	1,2-Дихлоретен	2
20	Діетиловий ефір	3
21	Етанол	3
22	Етилацетат	3
23	Етиленгліколь	2
24	Етилформіат	3
25	2-Етоксіетанол	2
26	Ізобутилацетат	3
27	Ізопропілацетат	3
28	Ксилол ²	2
29	Кумол	2
30	Метанол	2
31	Метилацетат	3
32	3-Метил-1-бутанол	3
33	Метилбутилкетон	2
34	Метилізобутилкетон	3
35	2-Метил-1-пропанол	3
36	<i>N</i> -Метилпіролідон	2
37	Метилциклогексан	2
38	Метилетилкетон	3
39	2-Метоксіетанол	2
40	Мурашина кислота	3
41	Нітрометан	2
42	Пентан	3
43	1-Пентанол	3
44	Піридин	2
45	1-Пропанол	3
46	2-Пропанол	3
47	Пропілацетат	3
48	Сульфолан	2
49	Тетрагідрофуран	2
50	Тетралін	2
51	Толуол	2
52	1,1,1-Трихлоретан	1

² Для розрахунку молекулярних дескрипторів використали саме *m*-ксилол, оскільки ксилол зазвичай містить 60% *m*-ксилолу, 14% *n*-ксилолу, 9% *o*-ксилолу та 17% етилбензолу [2].

53	1,1,2-Трихлоретен	2
54	Оцтова кислота	3
55	Формаїд	2
56	Хлорбензол	2
57	Хлороформ	2
58	Циклогексан	2
59	Тетрахлорметан	1

Таблиця 2.2. Розчинники, для яких відсутні достатні дані з токсичності [2]

Назва розчинника	
1,1-Диметоксиметан	Метилізопропілкетон
2,2-Диметоксипропан	Метилтетрагідрофуран
1,1-Діетоксипропан	Трихлороцтова кислота
Діізопропіловий ефір	Трифтороцтова кислота
Ізооктан	

2.2. Розрахунок молекулярних дескрипторів

За допомогою програмного пакету ChemOffice 2020 [22] для вказаних у таблицях 2.1–2.3 розчинників розраховано 18 молекулярних дескрипторів (хімічні, топологічні, фізичні та термодинамічні властивості) [23]:

- 1) молекулярна маса (molecular mass);
- 2) кількість акцепторів водневого зв'язку (number of HBond acceptors);
- 3) кількість донорів водневого зв'язку (number of HBond donors);
- 4) овальність (ovality):

$$O = A / (4 * \pi * ((3 * V) / 4 * \pi)^{2/3}), \quad (2.1)$$

де O – овальність, A – площа, V – об'єм;

- 5) молярна рефракція (molar refractivity):

$$A = \frac{4\pi}{3} N_A \alpha, \quad (2.2)$$

де A – молярна рефракція, N_A – стала Авогадро, α – середня поляризованість;

6) логарифмічний коефіцієнт розподілу між октанолом і водою ($\log P$), що є показником ліпофільності;

- 7) логарифмічний коефіцієнт розчинності у воді ($\log S$);

- 8) індекс Балабана (Balaban index):

$$J = \frac{q}{\mu+1} \sum (S_i S_j)^{-1/2}, \quad (2.3)$$

де q – кількість ребер у молекулярному графі, $\mu = q - n + 1$ – цикломатичне число молекулярного графа, n – число атомів в молекулярному графі, S_i – сума всіх елементів в i -му рядку (стовпчику) топологічної матриці відстаней;

9) кількість кластерів (cluster count) – кількість траєкторій заданої довжини в матриці відстаней;

10) молекулярний топологічний індекс (molecular topological index):

$$MTI = \sum_{i=1}^n E_i, \quad (2.4)$$

$$E = (A + D) * d, \quad (2.5)$$

де E – компоненти вектору, A – матриця суміжності, D – матриця відстані графа, d – вектор ступенів вершин графа.

11) площа полярної поверхні (polar surface area) – сума поверхні всіх полярних атомів (зазвичай Оксигену та Нітрогену), включаючи приєднані атоми Гідрогену;

12) атрибут форми (shape attribute) – вимірює розгалуження молекули;

13) коефіцієнт форми (shape coefficient):

$$I = \frac{D-R}{R}, \quad (2.6)$$

де D – максимальне значення діаметру між найбільш віддаленими атомами, R – мінімальне значення радіусу між найбільш центральними атомами;

14) сума ступенів (sum of degrees) – сума ступенів кожного атома (атомний ступінь – це кількість неводневих атомів, з якими він зв'язаний);

15) сума валентних ступенів (sum of valence degrees) – сума валентних ступенів кожного атома (валентний ступінь атома – це сума порядків його сусідніх зв'язків, включаючи водні);

16) загальна зв'язність (total connectivity) – зв'язність, що розглядається для усіх гетероатомів.

17) загальна валентна зв'язність (total valence connectivity) – валентна зв'язність, що розглядається для всіх гетероатомів;

18) індекс Вінера (Wiener index):

$$W = \frac{1}{2} \sum D_{ij}, \quad (2.7)$$

де D_{ij} – недіагональні елементи матриці відстаней.

Таблиця 2.3.* Значення 18 молекулярних дескрипторів для розчинників з відомим ступенем ризику/небезпеки відповідно до їх загрози здоров'ю людини

Назва розчинника	Номер молекулярного дескриптора																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Анізол	108,140	1	0	1,2486	3,3055	1,9218	-1,6508	4065	8	540	9,23	6,1250	0	16	26,0000	0,0722	0,0131	64
Ацетон	58,080	1	0	1,1844	1,6045	0,2004	-0,1336	213	4	84	17,07	2,2500	1	6	12,0000	0,5774	0,2041	9
Ацетонітрил	41,053	1	0	1,1459	1,1189	0,0373	-0,1284	47	3	44	23,79	1,3333	1	4	10,0000	0,7071	0,2236	4
Бензол	78,114	0	0	1,2040	2,6886	2,0193	-1,7812	972	6	264	0,00	4,1667	0	12	18,0000	0,1250	0,0370	27
1-Бутанол	74,123	1	1	1,2073	2,1857	0,8071	-0,3720	791	5	170	20,23	3,2000	1	8	12,0000	0,3536	0,1581	20
2-Бутанол	74,123	1	1	1,1704	2,1857	0,7217	-0,3208	712	5	158	20,23	3,2000	0	8	12,0000	0,4082	0,1826	18
Бутилацетат	116,160	1	0	1,3157	3,1490	1,5458	-1,1785	8753	8	606	26,30	6,1250	1	14	24,0000	0,1443	0,0295	79
<i>трет</i> -Бутилметиловий ефір	88,150	1	0	1,1644	2,6495	1,3197	-0,8691	1661	6	244	9,23	4,1667	0	10	14,0000	0,3536	0,2041	28
Гексан	86,178	0	0	1,2580	2,9602	3,6867	-2,7500	2076	6	320	0,00	4,1667	0	10	10,0000	0,2500	0,2500	35
Гептан	100,205	0	0	1,2934	3,4240	4,2396	-3,1641	4650	7	496	0,00	5,1429	1	12	12,0000	0,1768	0,1768	56
<i>N,N</i> -Диметилацетамід	87,122	1	0	1,2345	2,4370	-0,1794	-0,0134	1721	6	238	20,31	4,1667	0	10	18,0000	0,3333	0,0913	29
Диметилсульфоксид	78,129	1	0	1,1452	1,9444	-0,7147	0,3749	213	4	84	17,07	2,2500	1	6	8,6667	0,5774	0,5000	9
<i>N,N</i> -Диметилформамід	73,095	1	0	1,1991	1,9732	-0,3995	0,1983	712	5	148	20,31	3,2000	0	8	16,0000	0,4082	0,1054	18
1,2-Диметоксітан	90,122	2	0	1,2640	2,3388	-0,1725	0,2558	2076	6	268	18,46	4,1667	0	10	18,0000	0,2500	0,0833	35
1,4-Діоксан	88,106	2	0	1,1564	2,1614	-0,3648	0,0859	972	6	220	18,46	4,1667	0	12	12,0000	0,1250	0,0417	27
Дихлорметан	84,927	0	0	1,1213	1,6240	1,5004	-1,4336	47	3	24	0,00	1,3333	1	4	3,5556	0,7071	0,9091	4
1,2-Дихлоретан	98,954	0	0	1,1553	2,0878	1,7068	-1,5948	237	4	62	0,00	2,2500	0	6	5,5556	0,5000	0,6429	10
1,1-Дихлоретен	96,938	0	0	1,1514	2,0624	1,9157	-1,9385	213	4	60	0,00	2,2500	1	6	7,5556	0,5774	0,4546	9
1,2-Дихлоретен	96,938	0	0	1,1470	2,0624	2,2479	-2,1835	237	4	62	0,00	2,2500	0	6	7,5556	0,5000	0,4286	10
Діетиловий ефір	74,123	1	0	1,2490	2,1857	0,8656	-0,5100	791	5	176	9,23	3,2000	1	8	12,0000	0,3536	0,2041	20
Етанол	46,069	1	1	1,1496	1,2581	-0,0855	0,3046	47	3	40	20,23	1,3333	1	4	8,0000	0,7071	0,3162	4
Етилацетат	88,106	1	0	1,2569	2,2214	0,6532	-0,5017	1899	6	250	26,30	4,1667	1	10	20,0000	0,2887	0,0589	32
Етиленгліколь	62,068	2	2	1,1665	1,4112	-1,2190	1,0049	237	4	76	40,46	2,2500	0	6	14,0000	0,5000	0,1000	10
Етилформіат	74,079	1	0	1,2003	1,7576	0,4332	-0,2904	791	5	154	26,30	3,2000	1	8	18,0000	0,3536	0,0680	20
2-Етоксіетанол	90,122	2	1	1,2672	2,3388	-0,2679	0,4137	2076	6	266	29,46	4,1667	0	10	18,0000	0,2500	0,0645	35
Ізобутилацетат	116,160	1	0	1,2849	3,1490	1,3054	-1,0107	8209	8	570	26,30	6,1250	0	14	24,0000	0,1667	0,0340	74
Ізопропілацетат	102,133	1	0	1,2016	2,6852	1,1209	-0,8665	3992	7	374	26,30	5,1429	1	12	22,0000	0,2357	0,0481	48

Ксилол ³	106,168	0	0	1,2787	3,6162	2,8937	-2,5711	3882	8	552	0,00	6,1250	0	16	22,0000	0,0833	0,0278	61
Кумол	120,195	0	0	1,2577	4,0800	3,4236	-2,9726	7074	9	776	0,00	7,1111	0	18	24,0000	0,0589	0,0185	88
Метанол	32,042	1	1	1,1099	0,7943	-0,5134	0,6361	4	2	12	20,23	0,5000	0	2	6,0000	1,0000	0,4472	1
Метилацетат	74,079	1	0	1,2072	1,7576	0,2253	-0,1680	712	5	142	26,30	3,2000	0	8	18,0000	0,4082	0,0833	18
3-Метил-1-бутанол	88,150	1	1	1,1932	2,6495	1,3330	-0,7697	1899	6	266	20,23	4,1667	1	10	14,0000	0,2887	0,1291	32
Метилбутилкетон	100,161	1	0	1,2713	2,9959	1,5626	-1,1641	4321	7	429	17,07	5,1429	0	12	18,0000	0,2041	0,0722	52
Метилізобутилкетон	100,161	1	0	1,2296	2,9959	1,5356	-1,1465	3992	7	398	17,07	5,1429	1	12	18,0000	0,2357	0,0833	48
2-Метил-1-пропанол	74,123	1	1	1,1867	2,1857	0,5667	-0,2042	712	5	156	20,23	3,2000	0	8	12,0000	0,4082	0,1826	18
N-Метилпіролідон	99,133	1	0	1,2423	2,7234	-0,2037	-0,2788	1946	7	334	20,31	5,1429	0	14	22,0000	0,1179	0,0323	40
Метилциклогексан	98,189	0	0	1,1694	3,2466	3,8437	-3,2048	2050	7	392	0,00	5,1429	0	14	14,0000	0,1021	0,1021	42
Метилетилкетон	72,107	1	0	1,1857	2,0683	0,4566	-0,3374	712	5	158	17,07	3,2000	0	8	14,0000	0,4082	0,1443	18
2-Метоксіетанол	76,095	2	1	1,2157	1,8750	-0,6958	0,7457	791	5	152	29,46	3,2000	1	8	16,0000	0,3536	0,0913	20
Мурашина кислота	46,025	1	1	1,1055	0,8300	-0,3061	0,4091	47	3	32	37,30	1,3333	1	4	14,0000	0,7071	0,1054	4
Нітрометан	61,040	1	0	1,1546	1,2527	0,3278	-0,4563	213	4	66	51,81	2,2500	1	6	18,0000	0,5774	0,0745	9
Пентан	72,151	0	0	1,2219	2,4964	3,1337	-2,3358	791	5	192	0,00	3,2000	1	8	8,0000	0,3536	0,3536	20
1-Пентанол	88,150	1	1	1,2426	2,6495	1,3600	-0,7861	2076	6	288	20,23	4,1667	0	10	14,0000	0,2500	0,1118	35
Піридин	79,102	1	0	1,1881	2,4775	0,7674	-0,8595	972	6	253	12,36	4,1667	0	12	20,0000	0,1250	0,0287	27
1-Пропанол	60,096	1	1	1,1744	1,7219	0,2541	0,0420	237	4	90	20,23	2,2500	0	6	10,0000	0,5000	0,2236	10
2-Пропанол	60,096	1	1	1,1577	1,7219	0,3821	-0,0580	213	4	84	20,23	2,2500	1	6	10,0000	0,5774	0,2582	9
Пропілацетат	102,133	1	0	1,2853	2,6852	0,9928	-0,7644	4321	7	402	26,30	5,1429	0	12	22,0000	0,2041	0,0417	52
Сульфолан	120,166	2	0	1,1785	2,7277	0,3909	-1,2008	1898	7	336	34,14	5,1429	0	14	20,6667	0,1250	0,0510	39
Тетрагідрофуран	72,107	1	0	1,1632	2,0083	0,4968	-0,5119	375	5	144	9,23	3,2000	0	10	14,0000	0,1768	0,1021	15
Тетралін	132,206	0	0	1,2599	4,3664	3,4641	-3,2714	7954	10	960	0,00	8,1000	0	22	28,0000	0,0208	0,0069	109
Толуол	92,141	0	0	1,2410	3,1524	2,4565	-2,1816	2050	7	392	0,00	5,1429	0	14	20,0000	0,1021	0,0321	42
1,1,1-Трихлоретан	133,396	0	0	1,1410	2,5792	2,4686	-2,3969	633	5	88	0,00	3,2000	1	8	7,3333	0,5000	0,7289	16
1,1,2-Трихлоретен	131,380	0	0	1,1712	2,5538	2,6248	-2,6907	712	5	92	0,00	3,2000	0	8	9,3333	0,4082	0,4208	18
Оцтова кислота	60,052	1	1	1,1609	1,2938	-0,0861	0,1980	213	4	72	37,30	2,2500	1	6	16,0000	0,5774	0,0913	9
Формамід	45,041	1	1	1,1261	1,0456	-1,0465	0,5815	47	3	36	43,09	1,3333	1	4	12,0000	0,7071	0,1361	4
Хлорбензол	112,556	0	0	1,2239	3,1800	2,6413	-2,5203	2050	7	344	0,00	5,1429	0	14	19,7778	0,1021	0,0364	42
Хлороформ	119,369	0	0	1,1334	2,1154	2,0693	-2,0560	213	4	42	0,00	2,2500	1	6	5,3333	0,5774	0,8417	9
Циклогексан	84,162	0	0	1,1467	2,7828	3,3178	-2,8068	972	6	264	0,00	4,1667	0	12	12,0000	0,1250	0,1250	27

³ Для розрахунку молекулярних дескрипторів використали саме *m*-ксилол, оскільки ксилол зазвичай містить 60% *m*-ксилолу, 14% *n*-ксилолу, 9% *o*-ксилолу та 17% етилбензолу [2].

Тетрахлорметан	153,811	0	0	1,1401	2,6068	2,8581	-2,8769	633	5	64	0,00	3,2000	1	8	7,1111	0,5000	0,8265	16
----------------	---------	---	---	--------	--------	--------	---------	-----	---	----	------	--------	---	---	--------	--------	--------	----

Таблиця 2.4.* Значення 18 молекулярних дескрипторів для розчинників з невідомим ступенем ризику/небезпечності відповідно до їх загрози здоров'ю людини

Назва розчинника	Номер молекулярного дескриптора																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1,1-Диметоксиметан	76,095	2	0	1,2199	1,8750	-0,0342	0,1822	791	5	156	18,46	3,2000	1	8	16	0,3536	0,1179	20
2,2-Диметоксипропан	104,149	2	0	1,1802	2,8026	0,8478	-0,5374	3652	7	348	18,46	5,1429	1	12	20	0,2500	0,0833	44
1,1-Діетоксипропан	132,203	2	0	1,2882	3,7302	1,7035	-1,2031	13648	9	756	18,46	7,1111	1	16	24	0,1250	0,0417	96
Діізопропіловий ефір	102,177	1	0	1,2239	3,1133	1,8009	-1,2361	3992	7	408	9,23	5,1429	1	12	16	0,2357	0,1361	48
Ізооктан	114,232	0	0	1,1893	3,8878	4,7983	-3,5932	7322	8	584	0,00	6,1250	1	14	14	0,2041	0,2041	66
Метилізопропілкетон	100,161	1	0	1,2394	2,9959	1,5356	-1,1465	3992	7	398	17,07	5,1429	1	12	18	0,2357	0,0833	48
Метилтетрагідрофуран	86,134	1	0	1,1792	2,4721	0,9644	-0,8752	931	6	236	9,23	4,1667	0	12	16	0,1443	0,0833	26
Трихлороцтова кислота	163,378	1	1	1,1652	2,7680	1,0320	-1,1525	3493	7	198	37,3	5,1429	0	12	21,3333	0,2887	0,0665	42
Трифтороцтова кислота	114,023	4	1	1,1694	1,3403	0,2094	-0,2425	3493	7	198	37,3	5,1429	0	12	40	0,2887	0,0025	42

*Номер молекулярного дескриптору відповідає його номеру в порядку опису розрахованих характеристик у пункті 2.2.

2.3. Прийоми зниження розмірності даних

2.3.1. Розрахунок критеріїв Краскела-Уолліса

Критерій Краскела-Уолліса використовується для порівняння вибірок (груп, класів). Якщо розрахункове значення критерію більше критичного, то тестований параметр суттєво змінюється в залежності від класів розчинників, в інакшому випадку – немає статистично значимих міжкласових відмінностей для тестованого параметра [24].

У даній роботі за допомогою критерію Краскела-Уолліса визначали молекулярні дескриптори, значення яких найбільше впливає на віднесення розчинника до того чи іншого класу. Розрахунок критерію Краскела-Уолліса для 59 розчинників, що характеризуються 18 молекулярними дескрипторами, виконано із застосуванням програмного комплексу Matlab R2022b (individual license 11092188) [25] та наведено у таблиці 2.5.

Таблиця 2.5. Результати розрахунку критерію Краскела-Уолліса для 18 молекулярних дескрипторів

	Номер молекулярного дескриптору (див. пункт 2.2)																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
χ^2 *	5,32	11,06	4,42	4,94	0,37	4,37	6,23	2,03	1,35	3,26	10,3	1,37	9,24	1,81	5,75	3,22	5,92	2,02

*Критичне значення χ^2 при рівні значущості 5% та 10 %, відповідно, 5,99 та 4,60 (число ступенів свободи 2) [26].

Встановлено, що на класифікацію розчинників (за їх ступенем ризику/небезпечності відповідно до їх загрози здоров'ю людини) найбільший вплив становлять такі молекулярні дескриптори (їх значення суттєво змінюються в залежності від класу розчинника):

- 1) кількість акцепторів водневого зв'язку (number of HBond acceptors);
- 2) логарифмічний коефіцієнт розчинності у воді (logS);

3) площа полярної поверхні (polar surface area);

4) коефіцієнт форми (shape coefficient).

Також суттєво на класифікацію розчинників (за їх ступенем ризику/небезпечності відповідно до їх загрози здоров'ю людини) впливають такі параметри як сума валентних ступенів (sum of valence degrees) та загальна валентна зв'язність (total valence connectivity).

2.3.2. Метод головних компонент

Метод головних компонент (principal component analysis, PCA) – найпопулярніший метод зниження розмірності (стиснення) даних. Сутність методу полягає в поданні інформації меншою кількістю нових змінних через лінійну комбінацію вихідних даних, при цьому втрати від такого стиснення повинні бути зіставні з шумом даних [27].

Метод головних компонент виконано із застосуванням програмного комплексу Matlab R2022b (individual license 11092188) [25].

Інформативність розрахованих головних компонент для 59 розчинників, що характеризуються набором з 18 молекулярних дескрипторів, представлена у таблиці 2.6. Дана таблиця містить відсоток загальної дисперсії та відсоток кумулятивної дисперсії. Отже, одержані головні компоненти пояснюють 100% всієї варіації початкових показників, причому перша головна компонента вносить найбільший вклад та пояснює ~ 99,82% загальної дисперсії, інші три компоненти сумарно вносять суттєво менший вклад та пояснюють лише ~ 0,18%.

Таблиця 2.6. Оцінка рівня інформативності головних компонент для 59 розчинників, що характеризуються набором з 18 молекулярних дескрипторів

Головна компонента	Загальна дисперсія (%)	Кумулятивна дисперсія (%)
1	99,8238	99,8238
2	0,1482	99,9720
3	0,0248	99,9968
4	0,0032	100

Аналогічна ситуація спостерігається і в результаті розрахунку головних компонент для масиву даних з 18 молекулярних дескрипторів, що характеризують розчинники з невідомим ступенем/рівнем токсичності (таблиця 2.7). Одержані головні компоненти пояснюють 99,99% всієї варіації початкових показників, причому перша головна компонента вносить найбільший вклад та пояснює ~ 99,95% загальної дисперсії, інші три компоненти сумарно вносять суттєво менший вклад та пояснюють лише ~ 0,0496%.

Таблиця 2.7. Оцінка рівня інформативності головних компонент для 9 розчинників, що характеризуються набором з 18 молекулярних дескрипторів

Головна компонента	Загальна дисперсія (%)	Кумулятивна дисперсія (%)
1	99,9503	99,9503
2	0,0422	99,9925
3	0,0072	99,9997
4	0,0002	99,9999

Таким чином, у результаті застосування методу головних компонент вдалося суттєво знизити розмірність даних: кількість параметрів, що характеризують розчинники, з 18 зменшилась до 4. Значення головних компонент для 59 розчинників з відомим ступенем/рівнем токсичності та для 9 розчинників з невідомим ступенем/рівнем токсичності наведені, відповідно, у таблицях 2.8 та 2.9.

Таблиця 2.8. Значення головних компонент для 59 розчинників з відомим ступенем/рівнем токсичності

Назва розчинника	Головна компонента			
	1	2	3	4
Анізол	0,1936	-0,1251	-0,1214	-0,0153
Ацетон	0,0100	-0,0903	0,0681	0,0679
Ацетонітрил	0,0021	-0,0593	0,0662	0,1424
Бензол	0,0464	-0,2021	-0,0692	-0,0651
1-Бутанол	0,0376	-0,1185	0,0406	0,0747
2-Бутанол	0,0338	-0,1155	0,0510	0,0742
Бутилацетат	0,4164	0,3978	0,1733	0,0490
<i>трет</i> -Бутилметиловий ефір	0,0790	-0,0963	0,0405	-0,0298
Гексан	0,0989	-0,1271	-0,0396	-0,0961
Гептан	0,2215	0,0009	-0,0436	-0,1155
<i>N,N</i> -Диметилацетамід	0,0818	-0,0809	0,0554	0,0656
Диметилсульфоксид	0,0100	-0,0998	0,1226	0,0302
<i>N,N</i> -Диметилформаїд	0,0338	-0,1040	0,0621	0,0817
1,2-Диметоксіетан	0,0987	-0,0701	0,0487	0,0457
1,4-Діоксан	0,0462	-0,1573	0,0241	0,0456
Дихлорметан	0,0020	-0,0592	0,2017	-0,1392
1,2-Дихлоретан	0,0111	-0,0833	0,2030	-0,1554
1,1-Дихлоретен	0,0099	-0,0831	0,1985	-0,1493
1,2-Дихлоретен	0,0111	-0,0824	0,1977	-0,1493
Діетиловий ефір	0,0376	-0,1258	0,0280	-0,0113
Етанол	0,0021	-0,0572	0,0832	0,1004
Етилацетат	0,0903	-0,0714	0,0577	0,1131
Етиленгліколь	0,0110	-0,0788	0,1014	0,2497
Етилформаїт	0,0375	-0,1005	0,0651	0,1294
2-Етоксіетанол	0,0987	-0,0672	0,0560	0,1329
Ізобутилацетат	0,3905	0,3680	0,1808	0,0498
Ізопропілацетат	0,1899	0,0525	0,0873	0,0884
Ксилол ⁴	0,1850	-0,1615	-0,1603	-0,0849
Кумол	0,3370	-0,0079	-0,1820	-0,1016
Метанол	0,0000	-0,0256	0,0797	0,1172
Метилацетат	0,0337	-0,0975	0,0755	0,1290
3-Метил-1-бутанол	0,0903	-0,0894	0,0333	0,0581
Метилбутилкетон	0,2056	0,0339	0,0289	0,0209
Метилізобутилкетон	0,1899	0,0263	0,0456	0,0180
2-Метил-1-пропанол	0,0338	-0,1132	0,0537	0,0737
<i>N</i> -Метилпіролідон	0,0926	-0,1636	-0,0242	0,0689

⁴ Для розрахунку молекулярних дескрипторів використали саме *m*-ксилол, оскільки ксилол зазвичай містить 60% *m*-ксилолу, 14% *n*-ксилолу, 9% *o*-ксилолу та 17% етилбензолу [2].

Метилциклогексан	0,0978	-0,2153	-0,1057	-0,0941
Метилетилкетон	0,0338	-0,1148	0,0445	0,0563
2-Метоксіетанол	0,0375	-0,0989	0,0741	0,1477
Мурашина кислота	0,0020	-0,0474	0,1017	0,2453
Нітрометан	0,0098	-0,0699	0,1159	0,3462
Пентан	0,0377	-0,1433	-0,0029	-0,0856
1-Пентанол	0,0988	-0,0913	0,0172	0,0611
Піридин	0,0463	-0,1897	-0,0464	0,0333
1-Пропанол	0,0111	-0,0947	0,0685	0,0877
2-Пропанол	0,0100	-0,0910	0,0746	0,0859
Пропілацетат	0,2055	0,0634	0,0746	0,0905
Сульфолан	0,0903	-0,1811	0,0331	0,1443
Тетрагідрофуран	0,0178	-0,1427	0,0345	-0,0048
Тетралін	0,3791	-0,1042	-0,3304	-0,0871
Толуол	0,0978	-0,2123	-0,1218	-0,0732
1,1,1-Трихлоретан	0,0299	-0,0772	0,2912	-0,2146
1,1,2-Трихлоретен	0,0336	-0,0708	0,2864	-0,2085
Оцтова кислота	0,0099	-0,0767	0,0985	0,2297
Формамід	0,0020	-0,0512	0,0962	0,2922
Хлорбензол	0,0976	-0,1688	-0,0022	-0,1223
Хлороформ	0,0099	-0,0738	0,2835	-0,1958
Циклогексан	0,0464	-0,2051	-0,0531	-0,0859
Тетрахлорметан	0,0298	-0,0603	0,3788	-0,2565

Таблиця 2.9. Значення головних компонент для 9 розчинників з невідомим ступенем/рівнем токсичності

Назва розчинника	Головна компонента			
	1	2	3	4
1,1-Диметоксиметан	0,0445	0,3092	0,1715	0,1470
2,2-Диметоксипропан	0,2068	0,2993	0,0552	0,1312
1,1-Діетоксипропан	0,7744	-0,4655	-0,1647	0,0578
Діізопропіловий ефір	0,2262	0,3922	-0,0945	-0,0431
Ізооктан	0,4153	0,2447	-0,2630	-0,3230
Метилізопропілкетон	0,2262	0,3634	-0,0686	0,1828
Метилтетрагідрофуран	0,0524	0,4991	0,0388	0,0048
Трихлороцтова кислота	0,1975	-0,0144	0,7782	-0,5764
Трифтороцтова кислота	0,1976	-0,0560	0,5006	0,6971

Висновки до розділу 2

1. Встановлено, що на класифікацію розчинників (за їх ступенем ризику/небезпечності відповідно до їх загрози здоров'ю людини) найбільший вплив становлять такі молекулярні дескриптори (їх значення суттєво змінюються в залежності від класу розчинника): кількість акцепторів водневого зв'язку (number of HBond acceptors); логарифмічний коефіцієнт розчинності у воді (logS); площа полярної поверхні (polar surface area); коефіцієнт форми (shape coefficient).

2. У результаті застосування методу головних компонент вдалося суттєво знизити розмірність даних: кількість молекулярних дескрипторів, що характеризують розчинники, з 18 зменшилась до 4.

3. РЕЗУЛЬТАТИ ТА ЇХ ОБГОВОРЕННЯ

3.1. Навчання ймовірнісної нейронної мережі

Ймовірнісна нейронна мережа (Probabilistic Neural Network) – модифікація радіальних базисних нейронних мереж. Ймовірнісна нейронна мережа містить прихований шар нейронів з радіально-симетричною функцією активації, кожен із яких призначений для зберігання окремого еталонного зразка (відсутня потреба визначати число прихованих нейронів та вид функції активації):

$$F = e^{-\left[\frac{\sqrt{\sum_{i=1}^N (x_i - w_i)^2}}{2\delta}\right]^2} \quad (3.1)$$

де δ – відхилення функції, w_i – вагові коефіцієнти нейронів, x_i – числові характеристики досліджуваного об'єкту [28, 29].

Вихідний шар ймовірнісної мережі – конкуруючий шар, який підраховує ймовірність приналежності вхідного вектора до того чи іншого класу i , зрештою, зіставляє вектор з тим класом, ймовірність приналежності якого вище. Навчання ймовірнісної мережі передбачає попереднє проведення кластеризації для визначення центрів класів, найчастіше використовується алгоритм k -середніх [30, 31].

Основна проблема при реалізації ймовірнісної нейронної мережі полягає у визначенні параметра відхилення функції активації δ : значення параметра має бути досить великим, щоб перешкоджати перенавчанню, проте не настільки великим, щоб радіальна базова функція оголошувала однаково значущими всі значення входу. Вагові коефіцієнти першого шару формуються з використанням вхідних векторів з навчальної множини. Вагові коефіцієнти другого шару відповідають цільовим векторам навчальної вибірки.

Ймовірнісну нейронну мережу навчали на 59 розчинниках, для яких ступінь ризику/небезпечності відповідно до їх загрози здоров'ю людини відомі. Навчання проводили з використанням різного набору молекулярних

дескрипторів та головних компонент (за результатами розрахунку критерію Краскела-Уолліса та застосування методу головних компонент) та при різних значеннях параметру відхилення. Щоб контролювати якість навчання ймовірнісної нейронної мережі 59 розчинників поділили випадковим чином на навчальну вибірку (80%), тестову вибірку (10%) та контрольну (валідаційну) вибірку (10%). Склад тестової та контрольної вибірок наведено у таблиці 3.1, у навчальну вибірку увійшли усі розчинники, що не потрапили до цих вибірок.

Таблиця 3.1. Тестова та контрольна вибірки для контролю якості навчання ймовірнісної нейронної мережі

Тестова вибірка	Контрольна вибірка
1,2-Диметоксіетан	<i>N,N</i> -Диметилацетамід
1-Пентанол	Тетрахлорметан
1-Пропанол	2-Метил-1-пропанол
2-Пропанол	<i>N</i> -Метилпіролідон
1-Бутанол	3-Метил-1-бутанол
Гексан	Мурашина кислота

Ненадійність класифікації оцінювали як частку невірно класифікованих зразків тестової (контрольної) вибірки [32]:

$$P = \frac{n}{N} * 100\%, \quad (3.2)$$

де, n – число невірно класифікованих зразків тестової (контрольної) вибірки, N – загальне число зразків тестової (контрольної) вибірки.

У таблиці 3.2 наведено результати оцінки навчання ймовірнісної нейронної мережі з використанням усіх 18 розрахованих молекулярних дескрипторів. Максимальна надійність класифікації розчинників за їх ступенем ризику/небезпечності відповідно до їх загрози здоров'ю людини склала 50% при значеннях відхилення функції активації від 0,6 до 1,0.

Таблиця 3.2. Результати навчання ймовірнісної нейронної мережі з використанням усіх 18 розрахованих молекулярних дескрипторів

Параметр δ	<i>P</i> , %	
	Тестова вибірка	Контрольна вибірка
0,1	100,0	83,3
0,2	83,3	66,7
0,3	83,3	66,7
0,4	83,3	66,7
0,5	50,0	66,7
0,6	50,0	50,0
0,7	50,0	50,0
0,8	50,0	50,0
0,9	50,0	50,0
1,0	50,0	50,0

Таблиця 3.3 містить результати оцінки навчання ймовірнісної нейронної мережі з використанням 4 молекулярних дескрипторів, що за результатами розрахунку критерію Краскела-Уолліса мають найбільший вплив на віднесення розчинника до того чи іншого класу за їх ступенем ризику/небезпечності відповідно до їх загрози здоров'ю людини:

- 1) кількість акцепторів водневого зв'язку (number of HBond acceptors);
- 2) логарифмічний коефіцієнт розчинності у воді (logS);
- 3) площа полярної поверхні (polar surface area);
- 4) коефіцієнт форми (shape coefficient).

Зауважимо, що зниження розмірності даних до 4 молекулярних дескрипторів суттєво зменшило ненадійність класифікації.

Таблиця 3.3. Результати навчання ймовірнісної нейронної мережі з використанням 4 молекулярних дескрипторів (кількість акцепторів водневого зв'язку, логарифмічний коефіцієнт розчинності у воді, площа полярної поверхні, коефіцієнт форми)

Параметр δ	<i>P</i> , %	
	Тестова вибірка	Контрольна вибірка
0,1	16,7	33,3
0,2	16,7	33,3
0,3	16,7	33,3
0,4	16,7	33,3
0,5	16,7	33,3
0,6	16,7	33,3
0,7	16,7	50,0
0,8	16,7	50,0
0,9	16,7	33,3
1,0	16,7	33,3

У таблиці 3.4 наведено результати оцінки навчання ймовірнісної нейронної мережі з використанням 6 молекулярних дескрипторів, відбір яких базувався на результатах розрахунку критерію Краскела-Уолліса (до 4 молекулярних дескрипторів, що мають найбільший вплив на віднесення залишкових розчинників до того чи іншого класу за їх ступенем ризику/небезпечності відповідно до їх загрози здоров'ю людини, додали 2 дескриптори, що теж суттєво впливають на класифікацію залишкових розчинників):

- 1) кількість акцепторів водневого зв'язку (number of HBond acceptors);
- 2) логарифмічний коефіцієнт розчинності у воді (logS);
- 3) площа полярної поверхні (polar surface area);
- 4) коефіцієнт форми (shape coefficient);
- 5) сума валентних ступенів (sum of valence degrees);
- 6) загальна валентна зв'язність (total valence connectivity).

Додавання до досліджуваного масиву даних параметрів суми валентних ступенів та загальної валентної зв'язності дозволило досягти 100%

правильної класифікації залишкових розчинників як тестової вибірки, так і контрольної вибірки при усіх можливих значеннях відхилення функції активації, крім $\delta = 0,1$ (при $\delta = 0,1$ спостерігається невірне віднесення до 1 класу залишкового розчинника контрольної вибірки N-Метилпіролідону).

Таблиця 3.4. Результати навчання ймовірнісної нейронної мережі з використанням 6 молекулярних дескрипторів (кількість акцепторів водневого зв'язку, логарифмічний коефіцієнт розчинності у воді, площа полярної поверхні, коефіцієнт форми, сума валентних ступенів, загальна валентна зв'язність)

Параметр δ	<i>P, %</i>	
	Тестова вибірка	Контрольна вибірка
0,1	0,0	16,7
0,2	0,0	0,0
0,3	0,0	0,0
0,4	0,0	0,0
0,5	0,0	0,0
0,6	0,0	0,0
0,7	0,0	0,0
0,8	0,0	0,0
0,9	0,0	0,0
1,0	0,0	0,0

Таблиця 3.5 містить результати оцінки навчання ймовірнісної нейронної мережі з використанням 4 головних компонент, що описують 100% варіації початкових даних. Починаючи з $\delta = 0,3$ нейронна мережа усі розчинники тестової та контрольної вибірок відносить до 2 класу. При $\delta = 0,1$ та $\delta = 0,2$ спостерігаємо найменші помилки класифікації залишкових розчинників для даного випадку (16,7% для тестової вибірки та 50% для контрольної вибірки).

Таблиця 3.5. Результати навчання ймовірнісної нейронної мережі з використанням 4 головних компонент, що описують 100% варіації початкових даних

Параметр δ	<i>P, %</i>	
	Тестова вибірка	Контрольна вибірка
0,1	16,7	50,0
0,2	16,7	50,0
0,3	66,7	66,7
0,4	66,7	66,7
0,5	66,7	66,7
0,6	66,7	66,7
0,7	66,7	66,7
0,8	66,7	66,7
0,9	66,7	66,7
1,0	66,7	66,7

Зменшення кількості головних компонент до 3, 2 та 1 погіршило результати, усі розчинники при цьому відносяться до 2 класу (таблиці 3.6-3.8).

Таблиця 3.6. Результати навчання ймовірнісної нейронної мережі з використанням 3 головних компонент, що описують 99,9968% варіації початкових даних

Параметр δ	<i>P, %</i>	
	Тестова вибірка	Контрольна вибірка
0,1	16,7	50,0
0,2	66,7	66,7
0,3	66,7	66,7
0,4	66,7	66,7
0,5	66,7	66,7
0,6	66,7	66,7
0,7	66,7	66,7
0,8	66,7	66,7
0,9	66,7	66,7
1,0	66,7	66,7

Таблиця 3.7. Результати навчання ймовірнісної нейронної мережі з використанням 2 головних компонент, що описують 99,972% варіації початкових даних

Параметр δ	<i>P, %</i>	
	Тестова вибірка	Контрольна вибірка
0,1	66,7	66,7
0,2	66,7	66,7
0,3	66,7	66,7
0,4	66,7	66,7
0,5	66,7	66,7
0,6	66,7	66,7
0,7	66,7	66,7
0,8	66,7	66,7
0,9	66,7	66,7
1,0	66,7	66,7

Таблиця 3.8. Результати навчання ймовірнісної нейронної мережі з використанням 1 головної компоненти, що описує 99,8238% варіації початкових даних

Параметр δ	<i>P, %</i>	
	Тестова вибірка	Контрольна вибірка
0,1	66,7	66,7
0,2	66,7	66,7
0,3	66,7	66,7
0,4	66,7	66,7
0,5	66,7	66,7
0,6	66,7	66,7
0,7	66,7	66,7
0,8	66,7	66,7
0,9	66,7	66,7
1,0	66,7	66,7

Таким чином, найкраще навчання ймовірнісної нейронної мережі – з нульовою похибкою – забезпечує використання 6 молекулярних дескрипторів: кількість акцепторів водневого зв'язку, логарифмічний коефіцієнт розчинності у воді, площа полярної поверхні, коефіцієнт форми, сума валентних ступенів, загальна валентна зв'язність. Прогнозування

ступеня ризику/небезпечності залишкових розчинників у субстанціях, допоміжних речовинах і лікарських засобах буде здійснюватися саме на основі цих 6 параметрів.

3.2. Прогнозування ступеня ризику/небезпечності залишкових розчинників у субстанціях, допоміжних речовинах і лікарських засобах

Для прогнозування ступеня ризику/небезпечності залишкових розчинників у субстанціях, допоміжних речовинах і лікарських засобах використали ймовірнісну штучну нейронну мережу навчену на 59 розчинниках, що характеризуються 6 молекулярними дескрипторами: кількість акцепторів водневого зв'язку, логарифмічний коефіцієнт розчинності у воді, площа полярної поверхні, коефіцієнт форми, сума валентних ступенів, загальна валентна зв'язність.

Залишкові розчинники, для яких прогнозували ступінь ризику/небезпечності відповідно до їх загрози здоров'ю представлені у таблиці 2.2.

Також автором запропоновано валідаційну (контрольну) вибірку з 8 розчинників для перевірки запропонованого алгоритму прогнозування ступеня ризику/небезпечності залишкових розчинників (таблиця 3.9).

Таблиця 3.9. Розчинники, що запропоновані автором в якості валідаційної (контрольної) вибірки

Назва розчинника	
Пропіонітрил	Мезитилен
Етилбензол	Піролідон
Нітроетан	<i>o</i> -Ксилол
Трет-бутанол	<i>n</i> -Ксилол

Значення 6 молекулярних дескрипторів для 8 розчинників для перевірки запропонованого алгоритму прогнозування ступеня ризику/небезпечності залишкових розчинників були теж розраховані за

допомогою програмного пакету ChemOffice 2020 [3] та наведені у таблиці 3.10: кількість акцепторів водневого зв'язку (1), логарифмічний коефіцієнт розчинності у воді (2), площа полярної поверхні (3), коефіцієнт форми (4), сума валентних ступенів (5), загальна валентна зв'язність (6).

Таблиця 3.10. Значення 6 молекулярних дескрипторів для 8 розчинників валідаційної вибірки: кількість акцепторів водневого зв'язку (1), логарифмічний коефіцієнт розчинності у воді (2), площа полярної поверхні (3), коефіцієнт форми (4), сума валентних ступенів (5), загальна валентна зв'язність (6).

Назва розчинника	Номер молекулярного дескриптора					
	1	2	3	4	5	6
Пропіонітрил	1	-0,3319	23,79	0	12	0,1581
Етилбензол	0	-2,5695	0,00	0	22	0,0227
Нітроетан	1	-0,6706	51,81	0	20	0,0527
Трет-бутанол	1	-0,3857	20,23	1	12	0,2236
Мезитилен	0	-2,9535	0,00	0	24	0,0241
Піролідон	1	-0,1404	29,10	0	20	0,0361
<i>o</i> -Ксилол	0	-2,5644	0,00	0	22	0,0278
<i>n</i> -Ксилол	0	-2,5667	0,00	0	22	0,0278

Результати прогнозування ступеня ризику/небезпечності залишкових розчинників у субстанціях, допоміжних речовинах і лікарських засобах для 9 розчинників (таблиця 2.2) представлені у таблиці 3.11. Спостерігається однакова класифікація 9 розглянутих залишкових розчинників при значеннях відхилення функції активації від 0,1 до 1,0, крім трихлороцтової кислоти: при $\delta = 0,1$ вона класифікована як залишковий розчинник 1 класу, при $\delta = 0,2-1,0$ – як залишковий розчинник 2 класу).

Таблиця 3.11. Результати прогнозування ступеня ризику/небезпечності залишкових розчинників у субстанціях, допоміжних речовинах і лікарських засобах для 9 розчинників (таблиця 2.2)

Розчинник	Номер класу	
	$\delta = 0,1$	$\delta = 0,2-1,0$
1,1-Диметоксиметан	2	2
2,2-Диметоксипропан	2	2
1,1-Діетоксипропан	2	2
Діізопропіловий ефір	3	3
Ізооктан	2	2
Метилізопропілкетон	3	3
Метилтетрагідрофуран	3	3
Трихлороцтова кислота	1	2
Трифтороцтова кислота	1	1

Результати прогнозування ступеня ризику/небезпечності залишкових розчинників у субстанціях, допоміжних речовинах і лікарських засобах для 8 розчинників, що запропоновані автором в якості контрольної/валідаційної вибірки (таблиці 3.9, 3.10), представлені у таблиці 3.12. Спостерігається однакова класифікація усіх 8 розглянутих розчинників при значеннях відхилення функції активації від 0,1 до 1,0.

Таблиця 3.12. Результати прогнозування ступеня ризику/небезпечності залишкових розчинників у субстанціях, допоміжних речовинах і лікарських засобах для 8 розчинників, що запропоновані автором в якості контрольної/валідаційної вибірки (таблиця 3.9)

Розчинник	Номер класу	
	$\delta = 0,1$	$\delta = 0,2-1,0$
Пропіонітрил	2	2
Етилбензол	2	2
Нітроетан	2	2
Трет-бутанол	3	3
Мезитилен	2	2
Піролідон	2	2
<i>o</i> -Ксилол	2	2
<i>n</i> -Ксилол	2	2

Для уточнення приналежності до певного класу Трихлороцтової кислоти зробили наступне: ймовірнісну нейронну мережу навчили на 75 розчинниках (59 розчинників з відомим ступенем ризику/небезпечності відповідно до загрози здоров'ю людини, 8 розчинників тестової вибірки та 8 розчинників контрольної вибірки, для яких класова приналежність визначена безсумнівно), а Трихлороцтова кислота утворила тестову вибірку. Отримані такі результати: при $\delta = 0,1$ Трихлороцтова кислота віднесена до 1 класу залишкових розчинників, а при $\delta = 0,2-1,0$ Трихлороцтова кислота віднесена до 2 класу залишкових розчинників.

Оскільки 100% правильне навчання ймовірнісної нейронної мережі досягнуто при $\delta = 0,2-1,0$ (таблиця 2.13), то приймаємо, що Трихлороцтова кислота належить до 2 класу залишкових розчинників.

Зауважимо, що отримана класифікація залишкових розчинників на основі їх ступеня ризику/небезпечності відповідно до їх загрози здоров'ю людини відповідає їх хімічній природі, наприклад:

- 1) *o*-Ксилом та *n*-Ксилол в одній групі з *m*-Ксилом;
- 2) Піролідон в одній групі з *N*-Метилпіролідон;
- 3) Нітроетан в одній групі з Нітрометан;
- 4) Трет-бутанол в одній групі з 1-Бутанол та 2-Бутанол;
- 5) Діізопропіловий ефір в одній групі з іншими ефірами;
- 6) Мезитилен в одній групі з Толуол.

Висновки до розділу 3

1. Найкраще навчання ймовірнісної нейронної мережі – з нульовою похибкою – забезпечує використання 6 молекулярних дескрипторів: кількість акцепторів водневого зв'язку, логарифмічний коефіцієнт розчинності у воді, площа полярної поверхні, коефіцієнт форми, сума валентних ступенів, загальна валентна зв'язність.

2. Навчання нейронної мережі з використанням різної кількості головних компонент є не ефективним.

3. Отримана класифікація залишкових розчинників на основі їх ступеня ризику/небезпечності відповідно до їх загрози здоров'ю людини відповідає їх хімічній природі.

4. Застосування ймовірнісної нейронної мережі забезпечує надійну класифікацію залишкових розчинників при широкому діапазоні значень відхилення функції активації.

ВИСНОВКИ

1. Набір із шести молекулярних дескрипторів (кількість акцепторів водневого зв'язку, логарифмічний коефіцієнт розчинності у воді, площа полярної поверхні, коефіцієнт форми, сума валентних ступенів, загальна валентна зв'язність) є достатнім та оптимальним для правильної класифікації залишкових розчинників за їх ступенем ризику/небезпечності.

2. Застосування ймовірнісної нейронної мережі забезпечує надійну класифікацію залишкових розчинників за їх можливою загрозою здоров'ю людини при широкому діапазоні значень відхилення функції активації.

3. Запропонована процедура прогнозування ступеня небезпечності/ризиків залишкових розчинників у лікарських засобах може використовуватися при плануванні та створенні нових синтетичних лікарських засобів.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Impurities: Guideline for residual solvents. European Medicines Agency, Amsterdam, Netherlands, 2021, 51 p. URL: <https://www.ich.org/page/quality-guidelines> (date of access: 02.04.2023).
2. Державна Фармакопея України / Державне підприємство «Український науковий фармакопейний центр якості лікарських засобів». – 2-е вид. – Доповнення 2. Харків: Державне підприємство «Український науковий фармакопейний центр якості лікарських засобів», 2018. – 5.4. Залишкові розчинники. URL: <http://sphu.org/viddil-dfu> (date of access: 02.04.2023).
3. Zou L., Guo X., McElderry J.-D. Platform headspace gas chromatography method for high-throughput determination of residual solvents in pharmaceutical materials. *Journal of Pharmaceutical and Biomedical Analysis*. 2023. P. 115349. URL: <https://doi.org/10.1016/j.jpba.2023.115349> (date of access: 02.04.2023).
4. A generic gas chromatography method for determination of residual solvents in PET radiopharmaceuticals / I. Józszai et al. *Journal of Pharmaceutical and Biomedical Analysis*. 2022. Vol. 207. P. 114425. URL: <https://doi.org/10.1016/j.jpba.2021.114425> (date of access: 02.04.2023).
5. EL-Gindy A., Hadad G. M. Chemometrics in Pharmaceutical Analysis: An Introduction, Review, and Future Perspectives. *Journal of AOAC INTERNATIONAL*. 2012. Vol. 95, no. 3. P. 609–623. URL: https://doi.org/10.5740/jaoacint.sge_el-gindy (date of access: 02.04.2023).
6. Vu Dang H., Marin F. Editorial: Chemometrics-based Spectroscopy for Pharmaceutical and Biomedical Analysis. *Frontiers in Chemistry*. 2019. Vol. 7. URL: <https://doi.org/10.3389/fchem.2019.00153> (date of access: 02.04.2023).
7. The Use of Instrumental Technique and Chemometrics for Essential Oil Authentication: A Review / S. Syafri et al. *Results in Chemistry*. 2022. P. 100622. URL: <https://doi.org/10.1016/j.rechem.2022.100622> (date of access: 02.04.2023).

8. Mumali F. Artificial neural network-based decision support systems in manufacturing processes: A systematic literature review. *Computers & Industrial Engineering*. 2022. Vol. 165. P. 107964. URL: <https://doi.org/10.1016/j.cie.2022.107964> (date of access: 02.04.2023).
9. Guha S., Jana R. K., Sanyal M. K. Artificial Neural Network approaches for disaster management: A literature review (2010–2021). *International Journal of Disaster Risk Reduction*. 2022. P. 103276. URL: <https://doi.org/10.1016/j.ijdr.2022.103276> (date of access: 02.04.2023).
10. A review of artificial neural network based chemometrics applied in laser-induced breakdown spectroscopy analysis / L.-N. Li et al. *Spectrochimica Acta Part B: Atomic Spectroscopy*. 2021. Vol. 180. P. 106183. URL: <https://doi.org/10.1016/j.sab.2021.106183> (date of access: 02.04.2023).
11. Kaliuzhenko A., Pushkarova Y. Application of artificial neural networks for solving pharmaceutical issues. *Grail of Science*. 2023. No. 24. P. 766–769. URL: <https://doi.org/10.36074/grail-of-science.17.02.2023.143> (date of access: 02.04.2023).
12. Role of Artificial Neural Networks in Pharmaceutical Sciences / T. B. Teja et al. *Journal of Young Pharmacists*. 2022. Vol. 14, no. 1. P. 06–14. URL: <https://doi.org/10.5530/jyp.2022.14.2> (date of access: 02.04.2023).
13. Development of artificial intelligence powered apps and tools for clinical pharmacy services: a systematic review / F. Ranchon et al. *International Journal of Medical Informatics*. 2022. P. 104983. URL: <https://doi.org/10.1016/j.ijmedinf.2022.104983> (date of access: 02.04.2023).
14. Application of artificial neural networks to the prediction of antifungal activity of imidazole derivatives against *Candida albicans* / A. Badura et al. *Chemometrics and Intelligent Laboratory Systems*. 2022. Vol. 222. P. 104501. URL: <https://doi.org/10.1016/j.chemolab.2022.104501> (date of access: 02.04.2023).
15. Prediction of the antimicrobial activity of quaternary ammonium salts against *Staphylococcus aureus* using artificial neural networks / A. Badura et al.

Arabian Journal of Chemistry. 2021. Vol. 14, no. 7. P. 103233. URL: <https://doi.org/10.1016/j.arabjc.2021.103233> (date of access: 02.04.2023).

16. Prediction of biological activity of compounds containing a 1,3,5-triazinyl sulfonamide scaffold by artificial neural networks using simple molecular descriptors / E. Havránková et al. *Bioorganic Chemistry*. 2021. Vol. 107. P. 104565. URL: <https://doi.org/10.1016/j.bioorg.2020.104565> (date of access: 02.04.2023).

17. Pushkarova, Y., Zaitseva, G., & Al Saker, M. Prediction of Toxicity of Phenols Using Artificial Neural Networks. In 2022 12th International Conference on Advanced Computer Information Technologies (ACIT). 2022. P. 493-496. IEEE. URL: <https://ieeexplore.ieee.org/document/9913174> (date of access: 02.04.2023).

18. A novel toxicity prediction model for hydrazine compounds based on 1D–3D molecular descriptors / H.-y. Wang et al. *Computational Toxicology*. 2021. Vol. 18. P. 100169. URL: <https://doi.org/10.1016/j.comtox.2021.100169> (date of access: 02.04.2023).

19. Peng D., Picchioni F. Prediction of toxicity of Ionic Liquids based on GC-COSMO method. *Journal of Hazardous Materials*. 2020. Vol. 398. P. 122964. URL: <https://doi.org/10.1016/j.jhazmat.2020.122964> (date of access: 02.04.2023).

20. Wesoły M., Ciosek – Skibińska P. Comparison of various data analysis techniques applied for the classification of pharmaceutical samples by electronic tongue. *Sensors and Actuators B: Chemical*. 2018. Vol. 267. P. 570–580. URL: <https://doi.org/10.1016/j.snb.2018.04.050> (date of access: 02.04.2023).

21. Głowacz K., Wawrzyniak U. E., Ciosek-Skibińska P. Comparison of various data analysis techniques applied for the classification of oligopeptides and amino acids by voltammetric electronic tongue. *Sensors and Actuators B: Chemical*. 2021. Vol. 331. P. 129354. URL: <https://doi.org/10.1016/j.snb.2020.129354> (date of access: 02.04.2023).

22. The Prime Chemistry Portal. URL: <https://chemistrydocs.com/perkinelmer-chemoffice-2020-version-20-0/> (date of access: 02.04.2023).

23. Columbia University Libraries, Chem3D 17.0 User Guide. Copyright 1998-2017 PerkinElmer Informatics Inc., p. 203-232. URL: https://library.columbia.edu/content/dam/libraryweb/locations/dsc/Software%20Supportpages/ChemDraw_17_manual.pdf (date of access: 02.04.2023).

24. Ostertagová E., Ostertag O., Kováč J. Methodology and Application of the Kruskal-Wallis Test. *Applied Mechanics and Materials*. 2014. Vol. 611. P. 115–120. URL: <https://doi.org/10.4028/www.scientific.net/amm.611.115> (date of access: 02.04.2023).

25. Matlab for artificial intelligence. URL: <https://www.mathworks.com/products/matlab.html> (date of access: 02.04.2023).

26. Miller R. D., Miller J. C., Miller J. Statistics and Chemometrics for Analytical Chemistry. Pearson Education Australia, 2018. 296 p.

27. Principal Component Analysis / F. L. Gewers et al. *ACM Computing Surveys*. 2021. Vol. 54, no. 4. P. 1–34. URL: <https://doi.org/10.1145/3447755> (date of access: 02.04.2023).

28. Zeinali Y., Story B. A. Competitive probabilistic neural network. *Integrated Computer-Aided Engineering*. 2017. Vol. 24, no. 2. P. 105–118. URL: <https://doi.org/10.3233/ica-170540> (date of access: 02.04.2023).

29. Savchenko A. V., Belova N. S. Sequential analysis in fourier probabilistic neural networks. *Expert Systems with Applications*. 2022. P. 117885. URL: <https://doi.org/10.1016/j.eswa.2022.117885> (date of access: 02.04.2023).

30. Hoya T. Reducing the number of centers in a probabilistic neural network via applying the first neighbor means clustering algorithm. *Array*. 2022. Vol. 14. P. 100161. URL: <https://doi.org/10.1016/j.array.2022.100161> (date of access: 02.04.2023).

31. Ahmed M., Seraj R., Islam S. M. S. The k-means Algorithm: A Comprehensive Survey and Performance Evaluation. *Electronics*. 2020. Vol. 9,

no. 8. P. 1295. URL: <https://doi.org/10.3390/electronics9081295> (date of access: 02.04.2023).

32. Холін, Ю. В., Пушкарьова, Я. М., Пантелеймонов, А. В., & Некос, А. Н. Хемометричні методи в розв'язанні задач якісного хімічного аналізу та класифікації фізико-хімічних даних. Видавництво ХНУ імені В.Н. Каразіна, 2016. 184 с.

ДОДАТОК



Periodical scientific journal

№ 24 February 2023

GS 170223-212
dated 17.02.2023





Certificate of state registration of the print media KB24638-14578FP issued by the Ministry of Justice of Ukraine on 04.11.2020.

DOI 10.36074/grail-of-science/022023




CERTIFICATE OF PARTICIPATION AND PUBLICATION

Certificate provides at least a 0,3 ECTS credits to awarded participants for being involved.

Anastasiia Kaliuzhenko

participated in the V Correspondence International Scientific and Practical Conference
SCIENTIFIC RESEARCHES AND METHODS OF THEIR CARRYING OUT:
WORLD EXPERIENCE AND DOMESTIC REALITIES

held on February 17th, 2023 by | NGO European Scientific Platform (Winnytsia, Ukraine)
LLC International Centre Corporate Management (Vienna, Austria)

and published scientific paper

APPLICATION OF ARTIFICIAL NEURAL NETWORKS FOR SOLVING
PHARMACEUTICAL ISSUES

Euro Science Certificate № 22432 dated 06.01.2023

ISSN 2710-3056

UKRISTEI Certificate № 06 dated 09.01.2023

Head of the European Scientific Platform
Chairman of the Organizing committee
HOLDENBLAT MARIIA



Head of Community Outreach
LLC «International Centre Corporate Management»
RACHAEL APARO

