

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

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

**ВИПУСКНА КВАЛІФІКАЦІЙНА РОБОТА**

**на тему «ПРОГНОЗУВАННЯ БІОДОСТУПНОСТІ ЛІКАРСЬКИХ  
РЕЧОВИН ДЛЯ ПЕРОРАЛЬНОГО ЗАСТОСУВАННЯ НА ОСНОВІ ЇХ  
ХІМІЧНОЇ СТРУКТУРИ»**

Виконала: здобувачка вищої освіти 6 курсу  
групи 981А

напряму підготовки 22 «Охорона здоров'я»  
спеціальності 226 «Фармація, промислова  
фармація»

освітньої програми «Фармація»

Тимченко Інара Іграрівна

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

Рецензент: кандидат біологічних наук, доцент  
Кузнецова О.В.

КИЇВ-2025

## ЗМІСТ

ВСТУП .....	3
ОСНОВНА ЧАСТИНА.....	7
1. ЛІТЕРАТУРНИЙ ОГЛЯД .....	7
1.1 Біодоступність лікарських речовин: поняття та значення.....	7
1.2 Методи визначення біодоступності лікарських засобів .....	10
1.3 Штучні нейронні мережі: базові поняття та можливості .....	13
2. ЕКСПЕРИМЕНТАЛЬНА ЧАСТИНА .....	15
2.1 Матеріали та методи дослідження .....	15
2.2 Вибір інформативних дескрипторів.....	17
2.3 Навчання штучної нейронної мережі прямого поширення сигналу .....	21
2.4 Результати прогнозування біодоступності лікарських речовин із застосуванням нейронної мережі прямого поширення сигналу .....	32
ВИСНОВКИ.....	35
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	36
ДОДАТОК.....	40
SUMMARY .....	41

## ВСТУП

*Актуальність теми.* Прогнозування біодоступності лікарських речовин для перорального застосування є однією з ключових задач сучасної фармацевтичної науки, що знаходиться на перетині хімії, фармакології та біоінформатики. Ефективність фармакотерапії значною мірою залежить від біодоступності лікарських речовин, що визначає їх здатність досягати системного кровообігу в терапевтично значимій концентрації після перорального введення.

Біодоступність є комплексним показником, що залежить від багатьох факторів, таких як фізико-хімічні властивості молекули, її метаболізм в організмі, проникність через клітинні мембрани та стабільність в шлунково-кишковому тракті. Традиційні експериментальні методи визначення біодоступності, зокрема *in vitro* та *in vivo* дослідження, є трудомісткими, затратними за часом та фінансами, що обмежує їх використання на ранніх етапах розробки лікарських засобів.

Сучасні підходи, засновані на прогнозуванні біодоступності на основі хімічної структури молекули, відкривають можливості для значного пришвидшення розробки нових лікарських препаратів. Методи хемоінформатики та машинного навчання дозволяють створювати математичні моделі, що з високою точністю прогнозують біодоступність на основі структурних параметрів молекул. Це особливо важливо на етапі раннього відбору перспективних кандидатів серед численних хімічних сполук.

Нейронні мережі, завдяки своїм властивостям до навчання та адаптації до складних патернів у великих обсягах даних, є важливим інструментом у хемометрії. Вони можуть обробляти багатомірні та неструктуровані дані, що дає змогу розпізнавати складні взаємозв'язки між хімічними структурами молекул та їх біологічною активністю, включаючи біодоступність. Завдяки нейронним мережам також вдається враховувати не тільки хімічні

властивості молекул, а й інші важливі фактори, такі як метаболічні шляхи, механізми транспорту через біологічні мембрани та взаємодії з біологічними мішенями. Це підвищує точність прогнозів і дозволяє скоротити цикл розробки лікарських засобів, знижуючи витрати та прискорюючи вихід нових препаратів на ринок.

Тема є актуальною також у контексті розвитку персоналізованої медицини та вдосконалення технологій створення біоаналітичних моделей. Використання сучасних підходів до прогнозування біодоступності сприяє підвищенню ефективності процесу створення нових препаратів, зниженню витрат та скороченню часу виходу нових ліків на ринок.

Отже, розробка методології прогнозування біодоступності лікарських речовин для перорального застосування на основі їх хімічної структури із застосуванням штучної нейронної мережі прямого поширення сигналу є важливим та практично значущим напрямом, що відповідає сучасним вимогам фармацевтичної науки і потребам фармацевтичної індустрії.

*Мета і завдання дослідження.* Мета дослідження – оцінити ефективність алгоритму штучної нейронної мережі прямого поширення сигналу для прогнозування біодоступності лікарських речовин для перорального застосування за набором молекулярних дескрипторів.

Завдання дослідження:

1) розрахувати набір із 16 молекулярних дескрипторів для 145 лікарських речовин із застосуванням програмного забезпечення ChemOffice 2020 та визначити оптимальну їх комбінацію для прогнозування біодоступності молекул лікарських засобів;

2) реалізувати алгоритм штучної нейронної мережі прямого поширення сигналу та визначити її оптимальну архітектуру;

3) оцінити ефективність реалізованої нейронної мережі для вирішення задач прогнозування проникності лікарських речовин для перорального застосування за набором молекулярних дескрипторів.

*Методи дослідження:* множинна регресія, штучна нейронна мережа прямого поширення сигналу.

Програмне забезпечення Matlab R2024b (trial license) застосували для реалізації штучної нейронної мережі прямого поширення сигналу та множинної регресії між дескрипторами та біодоступністю.

*Новизна та значення одержаних результатів.* Встановлено, що для прогнозування біодоступності лікарських речовин для перорального застосування достатнім та інформативним є набір із семи молекулярних дескрипторів: молекулярна маса, логарифмічний коефіцієнт розподілу, кількість акцепторів водневого зв'язку, кількість донорів водневого зв'язку, індекс Балабана, молекулярний топологічний індекс, індекс Вінера. Практичне значення таких результатів полягає в тому, що встановлення ефективного набору молекулярних дескрипторів для прогнозування біодоступності лікарських речовин значно спрощує та прискорює процес оцінки потенціалу молекул на етапі їхнього попереднього відбору.

Показано, що нейронна мережа прямого поширення сигналу є надійним інструментом прогнозування біодоступності лікарських речовин на основі їх хімічної структури. Застосування методів машинного навчання для прогнозування біодоступності лікарських речовин є важливим кроком на шляху до створення більш ефективних, безпечних і економічно вигідних лікарських засобів. Це дозволяє значно зменшити витрати на лабораторні дослідження, оскільки відбір кандидатів для подальших експериментів можна здійснювати вже на основі точних математичних моделей, зменшуючи потребу в дорогих і трудомістких *in vitro* та *in vivo* тестах.

*Апробація результатів дослідження.* Результати цієї роботи були представлені на V Науково-практичній конференції з міжнародною участю «PLANTA+. НАУКА, ПРАКТИКА ТА ОСВІТА», присвяченій пам'яті доктора хімічних наук, професорки Ніни Павлівни Максютіної (до 100-річчя від дня народження), 28-29 січня 2025 р., м. Київ.

*Публікації.* Прийнята до друку стаття у фаховому журналі категорії Б «Сучасна медицина, фармація та психологічне здоров'я»: Yaroslava Pushkarova, Inara Tymchenko. Prediction of oral drug bioavailability based on chemical structure.

*Структура роботи:* 42 сторінки, 2 розділи, 1 додаток, список використаних джерел містить 31 публікацію.

# ОСНОВНА ЧАСТИНА

## 1. ЛІТЕРАТУРНИЙ ОГЛЯД

### 1.1 Біодоступність лікарських речовин: поняття та значення

Біологічна доступність лікарських речовин – це параметр, що характеризує ступінь та швидкість, з якою активна речовина лікарського засобу досягає системного кровообігу у незміненому вигляді після введення. Цей показник важливий для оцінки ефективності лікарського засобу [1].

Основні поняття, що стосуються оцінки біодоступності [2, 3]:

– абсолютна біодоступність – це показник, що відображає співвідношення кількості лікарської речовини, яка досягає системного кровообігу у незміненому вигляді після введення будь-яким шляхом, крім внутрішньовенного, (наприклад, пероральним), до кількості, введеної внутрішньовенно; вимірюється у відсотках;

– відносна біодоступність – це порівняння біодоступності двох різних форм одного і того самого лікарського засобу (наприклад, таблетована форма vs рідка форма);

– швидкість біодоступності – це характеристика, яка відображає, як швидко лікарська речовина досягає системного кровообігу після введення.

Швидкість біодоступності визначається кількома ключовими параметрами [4]:

– час досягнення максимальної концентрації – це проміжок часу від моменту введення лікарського засобу до досягнення максимальної концентрації активної речовини у плазмі крові; менше значення часу досягнення максимальної концентрації свідчить про швидшу абсорбцію речовини;

– максимальна концентрація – найвища концентрація лікарської речовини, яка досягається у плазмі крові.

Біодоступність – це критично важливий параметр у фармакології, який характеризує ефективність лікарського засобу. Значення біодоступності полягає у визначенні того, наскільки повно і швидко активна речовина досягає місця своєї дії. Це впливає на терапевтичний ефект препарату, його безпеку, дозування та вибір форми випуску.

Розглянемо роль та значення біологічної доступності у фармації детальніше [5, 6, 7].

*Оцінка ефективності препарату:*

– забезпечення терапевтичної дії – біодоступність показує, яка частка введеної дози досягає системного кровообігу у незміненому вигляді. Якщо цей показник низький, навіть велика доза препарату може виявитися неефективною;

– контроль фармакокінетичних параметрів: завдяки біодоступності визначають такі параметри, як площа під фармакокінетичною кривою «концентрація–час», максимальна концентрація в плазмі, і час її досягнення. Ці показники дозволяють оцінити, наскільки препарат відповідає очікуванням;

*Аналіз можливих способів введення препарату* – дослідження того, як різні методи введення (пероральний, ін'єкційний, трансдермальний тощо) впливають на біодоступність лікарського засобу та його здатність досягати потрібної терапевтичної концентрації в організмі.

*Оптимізація лікарських препаратів і біодоступність* – це важливе поєднання при розробці нових лікарських форм, оскільки біодоступність визначає, скільки активної речовини досягає системного кровообігу і таким чином визначає ефективність лікування.

*Можливі шляхи оптимізації лікарських препаратів і підвищення біодоступності:*

– мікронізація частинок – це процес подрібнення частинок активної речовини до дуже малих розмірів, що дозволяє збільшити її розчинність і, відповідно, біодоступність. Маленькі частинки мають більшу поверхню для



контакту з рідиною в шлунково-кишковому тракті, що полегшує їх абсорбцію в кров. Мікронізація часто використовується для ліків з низькою водорозчинністю;

– використання носіїв, таких як ліпосоми, оскільки ліпосоми – це мікроскопічні бульбашки, що складаються з фосфоліпідних мембран, які можуть інкапсулювати лікарську речовину. Цей метод дозволяє захистити активний компонент від метаболізму в шлунково-кишковому тракті та забезпечити його поступове вивільнення, підвищуючи біодоступність і тривалість дії. Ліпосоми також можуть допомогти направити ліки до специфічних тканин чи органів, покращуючи терапевтичний ефект;

– *створення пролонгованих форм* – пролонговані форми лікарських засобів (наприклад, таблетки з поступовим вивільненням) забезпечують повільне і контрольоване вивільнення активного компонента протягом тривалого часу. Це дозволяє підтримувати стабільну концентрацію препарату в крові, зменшуючи потребу в частих дозах і покращуючи біодоступність за рахунок кращої абсорбції.

#### *Індивідуалізація терапії:*

– вибір ліків для конкретних пацієнтів – індивідуальний підхід до вибору лікарських засобів передбачає, що лікар враховує специфічні особливості пацієнта перед тим, як призначити лікування (вік пацієнта, генетичні особливості, патології та супутні захворювання);

– оцінка взаємодії з їжею та іншими препаратами – прийом їжі може змінювати абсорбцію ліків, впливати на їх метаболізм і виведення з організму; препарати можуть взаємодіяти між собою, що може вплинути на їх ефективність або спричинити побічні ефекти.

#### *Безпека і токсичність*

– уникнення токсичних концентрацій – висока біодоступність може спричинити різкі піки концентрації препарату в крові, що підвищує ризик побічних ефектів. Тому препарати з вузьким терапевтичним індексом потребують ретельного контролю дозування;

– запобігання недостатньому ефекту – занадто низька біодоступність може бути причиною того, що препарат не досягає терапевтичної концентрації у крові, і пацієнт не отримує бажаного лікувального ефекту.

Також, значення біодоступності використовується для визначення еквівалентності між оригінальним препаратом і його генеричними версіями. Якщо біодоступність генерика значно відрізняється від оригіналу, це може вплинути на його ефективність або безпеку.

Отже, для розробки ефективних лікарських форм необхідно оптимізувати біодоступність таким чином, щоб максимальна кількість активної речовини досягала системного кровообігу, що дозволить забезпечити високу ефективність лікування та зменшити побічні ефекти. Це вимагає комплексного підходу, який включає вибір правильного шляху введення, покращення розчинності та лікарських засобів.

## **1.2 Методи визначення біодоступності лікарських засобів**

Експериментальні методи визначення біодоступності лікарських засобів спрямовані на оцінку того, яка кількість активної речовини досягає системного кровообігу і стає доступною для біологічного впливу. Вибір методу залежить від фізико-хімічних властивостей препарату, шляху введення та цільового призначення.

Основні підходи до визначення біодоступності лікарських засобів включають *in vivo* та *in vitro* методи [8-10].

*In vivo методи* – це експериментальні підходи, які використовуються для визначення біодоступності лікарських засобів безпосередньо в живих організмах. Вони дозволяють оцінити, як лікарська речовина взаємодіє з біологічними системами, і враховують усі фізіологічні та біохімічні процеси, які впливають на абсорбцію, розподіл, метаболізм та екскрецію препарату. Приклади: дослідження фармакокінетики препарату у людини для визначення параметрів, таких як максимальна концентрація та площа під

кривою концентрація-час, вимірювання рівня лікарської речовини в крові, тканинах або сечі після введення.

*In vitro методи* – це експериментальні підходи, що проводяться поза живим організмом у контрольованих лабораторних умовах. Вони моделюють окремі процеси, які відбуваються в організмі, і дозволяють оцінити ключові характеристики лікарських засобів на етапі розробки та скринінгу.

Основні аспекти *in vitro методів*:

– методи розчинності та дисоціації – дослідження швидкості розчинення лікарської речовини з різним рН (імітація середовища шлунка, кишечника);

– клітинні моделі – використання моношарів клітин (наприклад, Caco-2 клітини) для оцінки проникності речовини через епітелій кишечника;

– трансдермальні моделі – дослідження проникності через штучні або природні мембрани для лікарських форм, призначених для зовнішнього застосування;

– дослідження стабільності – оцінка стійкості лікарської речовини до ферментів шлунка, кишкових соків або мікрофлори;

– вивчення взаємодії з біомолекулами – оцінка зв'язування лікарської речовини з білками плазми крові, аналіз метаболізму за допомогою ферментів.

*In silico методи* – це комп'ютерне моделювання та математичні підходи, що дозволяють прогнозувати біодоступність лікарських засобів, ґрунтуючись на їх фізико-хімічних властивостях, структурі молекули та великих масивах даних. Ці методи є потужним інструментом у сучасній фармацевтичній розробці, адже дозволяють зменшити витрати часу та ресурсів у порівнянні з *in vitro* та *in vivo* підходами [11].

Ключові аспекти *in silico* методів [12-14]:

– ADME-аналіз (Absorption, Distribution, Metabolism, Excretion) – використовує математичні моделі та симуляції для прогнозування шляхів

абсорбції, розподілу, метаболізму та екскреції лікарської речовини (наприклад, програмні пакети Simcyp Simulator, GastroPlus);

– моделі молекулярної динаміки – досліджують взаємодію молекули лікарської речовини з біомолекулами (ферментами, білками-мішенями, рецепторами) за допомогою програмних пакетів AMBER, GROMACS, AutoDock тощо;

– машинне навчання (Machine Learning) – алгоритми машинного навчання аналізують великі бази даних експериментальних результатів і створюють прогностичні моделі біодоступності; Моделі, побудовані за допомогою штучних нейронних мереж, регресійних аналізів або методів кластеризації, допомагають виявити закономірності між структурою молекули та її фармакокінетичними параметрами (наприклад, платформи KNIME, TensorFlow, Scikit-learn);

– QSAR-аналіз (Quantitative Structure-Activity Relationship) – метод встановлює зв'язок між структурою молекули та її біологічною активністю чи фармакокінетичними характеристиками (наприклад, програмне забезпечення ChemAxon, ACD/Labs, MOE);

– віртуальні скринінги (Virtual Screening) – оцінюють потенційну біодоступність тисяч молекул, обираючи найперспективніші для подальших досліджень;

– моделювання фармакокінетики та фармакодинаміки (PBPK/PD) – PBPK-моделі (фізіологічно обґрунтовані фармакокінетичні моделі) використовуються для прогнозування концентрації лікарської речовини в крові та тканинах; допомагають оцінити варіативність між пацієнтами залежно від віку, статі, стану здоров'я.

In silico методи є невід'ємною частиною сучасного фармацевтичного дослідження. Вони дозволяють прискорити процес розробки лікарських засобів, оптимізувати витрати та виявляти перспективні молекули на етапі доклінічного скринінгу.

### **1.3 Штучні нейронні мережі: базові поняття та можливості**

Штучні нейронні мережі – це інструменти математичного моделювання, натхнені роботою людського мозку, які здатні навчатися, аналізувати дані, знаходити закономірності та приймати рішення. Їх популярність пов'язана зі здатністю адаптуватися до складних завдань, що вимагають значного обсягу обчислень, таких як розпізнавання образів, обробка мови, медична діагностика тощо.

У штучній нейронній мережі нейрон є основною обчислювальною одиницею. Нейрони в штучних мережах працюють за принципом аналогії з біологічними нейронами, де кожен нейрон отримує вхідну інформацію, обробляє її та передає результат наступному нейрону через синапс, що представляється вагою у штучних мережах. Він виконує функцію перетворення вхідного сигналу в результат, застосовуючи до нього вагові коефіцієнти та активаційну функцію. Вхідні дані проходять через ваги, які оновлюються під час навчання, щоб мінімізувати похибки в результатах. Нейрони об'єднуються у шари відповідно до свого призначення. Вхідний шар нейронів приймає дані, які надходять у вигляді чисел (наприклад, інтенсивність пікселів у зображенні або значення параметрів). Приховані шари виконують обробку даних. Кількість шарів залежить від складності поставленої задачі. Вихідний шар формує результат, наприклад, визначення класу об'єкта або прогнозування числа [15].

Активаційні функції дозволяють мережі моделювати складні нелінійні залежності у даних.

Переваги штучних нейронних мереж [16]:

- здатність обробляти складні й великі масиви даних;
- автоматичне виявлення закономірностей без ручного втручання;
- універсальність: можуть використовуватися в багатьох сферах;
- можливість роботи з нелінійними залежностями в даних.

Можливості штучних нейронних мереж [17, 18]:

– класифікація та регресія – нейронні мережі використовуються для розв’язання задач класифікації (визначення категорії чи класу) або регресії (прогнозування числових значень);

– обробка зображень – нейронні мережі дозволяють розпізнавати об’єкти на зображеннях, сегментувати медичні знімки або навіть відновлювати пошкоджені частини фото;

– обробка тексту та природної мови (Natural Language Processing);

– аналіз великих даних – нейронні мережі виявляють закономірності в масивних наборах даних.

Штучні нейронні мережі є одним із найпотужніших інструментів сучасних технологій. Вони здатні розв’язувати задачі, які раніше були недоступними для традиційних алгоритмів, і відкривають нові горизонти в науці, медицині, фармації та багатьох інших галузях.

## 2. ЕКСПЕРИМЕНТАЛЬНА ЧАСТИНА

### 2.1 Матеріали та методи дослідження

Програмне забезпечення ChemOffice 2020 [19] застосували для розрахунку набору молекулярних дескрипторів. Програмне забезпечення Matlab R2024b (trial license) [20] застосували для реалізації штучної нейронної мережі прямого поширення сигналу та множинної регресії між дескрипторами та біодоступністю.

Нейронна мережа прямого поширення — це базовий тип штучної нейронної мережі, в якій інформація передається лише в одному напрямку: від вхідного шару через приховані шари до вихідного шару. Сигнал проходить через мережу без зворотного зв'язку, що робить її простою у структурі та реалізації [21].

Параметри навчання для реалізованої нейронної мережі прямого поширення сигналу: метод ініціалізації вагових коефіцієнтів – алгоритм Нгуена-Відроу, метод навчання – Левенберга-Марквардта, функція активації для прихованого шару нейронів – тангенсна, функція активації для вихідного шару нейронів – лінійна [22].

Множинна регресія – це статистичний метод аналізу, який використовується для моделювання залежності між однією залежною змінною (відгуком) і декількома незалежними змінними (предикторами). Вона є узагальненням простої лінійної регресії, що враховує кілька факторів одночасно [23].

Досліджуваний масив даних містив 145 лікарських речовин для перорального застосування. Для кожної лікарської речовини відомо експериментальне значення біодоступності ( $y$  %) з літературних джерел [24, 25].

Для кожної молекули лікарської речовини розраховали 16 молекулярних дескрипторів на основі їх хімічної структури:

– параметр 1 – молекулярна маса;

- параметр 2 – кількість акцепторів водневого зв'язку;
- параметр 3 – кількість донорів водневого зв'язку;
- параметр 4 – овальність (ovality):

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

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

- параметр 5 – молярна рефракція;
- параметр 6 – логарифмічний коефіцієнт розподілу між октанолом і водою;
- параметр 7 – логарифмічний коефіцієнт розчинності у воді;
- параметр 8 – індекс Балабана:

$$J = \frac{q}{\mu+1} \sum (S_i S_j)^{-1/2}, \quad (2.2)$$

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;

- параметр 9 – кількість кластерів (cluster count) – кількість траєкторій заданої довжини в матриці відстаней;
- параметр 10 – молекулярний топологічний індекс:

$$MTI = \sum_{i=1}^n E_i, \quad (2.3)$$

$$E = (A + D) * d, \quad (2.4)$$

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

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

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

- параметр 13 – коефіцієнт форми (shape coefficient):

$$I = \frac{D-R}{R}, \quad (2.5)$$



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

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

– параметр 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);

– параметр 16 – індекс Вінера:

$$W = \frac{1}{2} \sum D_{ij}, \quad (2.6)$$

де  $D_{ij}$  – недіагональні елементи матриці відстаней.

## 2.2 Вибір інформативних дескрипторів

Відправною точкою для визначення необхідного та інформативного набору молекулярних дескрипторів для прогнозування біодоступності лікарських речовин для перорального застосування був набір із чотирьох параметрів відповідно до «правила п'яти» Ліпінського: молекулярна маса, логарифмічний коефіцієнт розподілу між октанолом і водою, кількість акцепторів водневого зв'язку та кількість донорів водневого зв'язку [26, 27]. До вказаного початкового набору дескрипторів по черзі додавали інші дескриптори та оцінювали їхній вплив на прогностичну модель.

Якість прогностичної моделі оцінювали за допомогою множинних коефіцієнтів детермінації для регресії між дескрипторами (незалежними змінними) та біодоступністю (залежною змінною чи відгуком).

У таблиці 1.2 представлені значення множинних коефіцієнтів детермінації для регресії між біодоступністю та різними наборами дескрипторів.

**Таблиця 1.2.** Значення множинних коефіцієнтів детермінації для регресії між біодоступністю та різними наборами дескрипторів

Набір дескрипторів	Коефіцієнт детермінації (R)
молекулярна маса, логарифмічний коефіцієнт розподілу, кількість акцепторів водневого зв'язку та кількість донорів водневого зв'язку (початковий набір дескрипторів)	0,2306
(молекулярна маса, логарифмічний коефіцієнт розподілу, кількість акцепторів водневого зв'язку та кількість донорів водневого зв'язку) + овальність	0,2312
(молекулярна маса, логарифмічний коефіцієнт розподілу, кількість акцепторів водневого зв'язку та кількість донорів водневого зв'язку) + молярна рефракція	0,2502
(молекулярна маса, логарифмічний коефіцієнт розподілу, кількість акцепторів водневого зв'язку та кількість донорів водневого зв'язку) + логарифмічний коефіцієнт розчинності у воді	0,2664
(молекулярна маса, логарифмічний коефіцієнт розподілу, кількість акцепторів водневого зв'язку та кількість донорів водневого зв'язку) + індекс Балабана	<b>0,3195</b>
(молекулярна маса, логарифмічний коефіцієнт розподілу, кількість акцепторів водневого зв'язку та кількість донорів водневого зв'язку) + кількість кластерів	0,2401
(молекулярна маса, логарифмічний коефіцієнт розподілу, кількість акцепторів водневого зв'язку та кількість донорів водневого зв'язку) + молекулярний топологічний індекс	<b>0,3312</b>
(молекулярна маса, логарифмічний коефіцієнт розподілу, кількість акцепторів водневого зв'язку та кількість донорів водневого зв'язку) + площа полярної поверхні	0,2399

(молекулярна маса, логарифмічний коефіцієнт розподілу, кількість акцепторів водневого зв'язку та кількість донорів водневого зв'язку) + атрибут форми	0,2403
(молекулярна маса, логарифмічний коефіцієнт розподілу, кількість акцепторів водневого зв'язку та кількість донорів водневого зв'язку) + коефіцієнт форми	0,2315
(молекулярна маса, логарифмічний коефіцієнт розподілу, кількість акцепторів водневого зв'язку та кількість донорів водневого зв'язку) + сума ступенів	0,2508
(молекулярна маса, логарифмічний коефіцієнт розподілу, кількість акцепторів водневого зв'язку та кількість донорів водневого зв'язку) + сума валентних ступенів	0,2336
(молекулярна маса, логарифмічний коефіцієнт розподілу, кількість акцепторів водневого зв'язку та кількість донорів водневого зв'язку) + індекс Вінера	<b>0,2967</b>
(молекулярна маса, логарифмічний коефіцієнт розподілу, кількість акцепторів водневого зв'язку та кількість донорів водневого зв'язку) + індекс Балабана + молекулярний топологічний індекс	<b>0,3321</b>
(молекулярна маса, логарифмічний коефіцієнт розподілу, кількість акцепторів водневого зв'язку та кількість донорів водневого зв'язку) + індекс Балабана + молекулярний топологічний індекс + індекс Вінера	<b>0,3701</b>
(молекулярна маса, логарифмічний коефіцієнт розподілу, кількість акцепторів водневого зв'язку та кількість донорів водневого зв'язку) + індекс Балабана + молекулярний топологічний індекс + індекс Вінера + логарифмічний коефіцієнт розчинності у воді	0,3729

Результати, наведені у таблиці 1.2, свідчать, що додавання індексу Балабана, молекулярно-топологічного індексу та індексу Вінера до початкового набору дескрипторів справді покращує регресійну модель – спостерігається суттєве збільшення множинного коефіцієнта детермінації.

Додавання до початкового набору дескрипторів таких параметрів як овальність, молярна рефракція, логарифмічний коефіцієнт розчинності у воді, кількість кластерів, площа полярної поверхні, атрибут форми, коефіцієнт форми, сума ступенів, сума валентних ступенів не покращує прогностичну модель – спостерігається зовсім незначна зміна у значенні множинного коефіцієнта детермінації у порівнянні з його значенням для початкового набору дескрипторів.

Показано, що набір із семи дескрипторів (молекулярна маса, логарифмічний коефіцієнт розподілу, кількість акцепторів водневого зв'язку, кількість донорів водневого зв'язку, індекс Балабана, молекулярний топологічний індекс, індекс Вінера) найкраще описує варіабельність біодоступності ( $R = 0,3701$ ). Рівняння для моделі множинної регресії з використанням семи вищезгаданих параметрів має такий вигляд: біодоступність =  $48,8405 + 0,05167 \cdot \text{молекулярна маса} + 0,08941 \cdot \text{кількість акцепторів водневих зв'язків} - 1,8777 \cdot \text{кількість донорів водневих зв'язків} + 4,5299 \cdot \text{логарифмічний коефіцієнт розподілу} - 0,00001634 \cdot \text{індекс Балабана} - 0,0145 \cdot \text{молекулярно-топологічний індекс} + 0,09957 \cdot \text{індекс Вінера}$ .

Додавання логарифмічного коефіцієнта розчинності у воді як восьмого дескриптора трохи збільшило множинний коефіцієнт детермінації, але незначно (від 0,3701 до 0,3729, різниця складає менше 1%).

Отже, оптимальний набір дескрипторів для прогнозування біодоступності лікарських речовин для перорального застосування включає сім параметрів: молекулярна маса, логарифмічний коефіцієнт розподілу, кількість акцепторів водневого зв'язку, кількість донорів водневого зв'язку, індекс Балабана, молекулярний топологічний індекс, індекс Вінера.

### **2.3 Навчання штучної нейронної мережі прямого поширення сигналу**

Досліджуваний набір із 145 лікарських речовин випадковим чином розділили на навчальну вибірку (130 речовин), тестову вибірку (10 речовин) та валідаційну вибірку (5 речовин).

Навчальна вибірка – набір даних, що використовується для навчання моделі, тобто для підбору параметрів моделі. Валідаційна вибірка – набір даних, який використовується для оцінки ефективності моделі під час навчання та її налаштування. Після початкового навчання модель перевіряється на валідаційній вибірці, тобто оцінюється ефективність модулі на нових даних. Тестова вибірка – набір даних, що використовується для кінцевої продуктивності моделі після завершення її навчання та налаштування. Тестова вибірка дає уявлення про те, наскільки добре модель буде працювати на повністю нових даних [28, 29].

Перелік досліджуваних 145, їх розподіл по вибіркам, а також значення семи параметрів, що обрані як інформативні та достатні, представлено у таблиці 2.2.

**Таблиця 2.2.** Список лікарських речовин, їх розподіл по вибіркам і розраховані значення семи молекулярних дескрипторів

№	Лікарська речовина	Вибірка	Дескриптор*						
			1	2	3	4	5	6	7
1	Ацетамінефон	Навчальна	151,17	2	2	0,441	19887	1227	166
2	Ацикловір	Навчальна	225,21	7	3	-1,622	81102	3053	453
3	Алендронат	Навчальна	249,10	4	6	-2,802	101097	1756	281
4	Алопуринол	Навчальна	136,11	4	2	-0,485	7659	777	105
5	Амілорид	Навчальна	229,63	7	4	1,085	81960	2405	368
6	Аміноглутетимід	Навчальна	232,28	3	2	1,308	97697	3596	484
7	Амлодипін	Навчальна	408,88	5	2	1,840	981585	12808	1836
8	Амоксицилін	Навчальна	365,40	5	4	0,045	518482	10602	1551
9	Ампіцилін	Навчальна	349,41	4	3	0,454	421931	9609	1365
10	Аспірин	Навчальна	180,16	2	1	1,443	41162	1722	246
11	Атенолол	Навчальна	266,34	4	3	0,460	317560	6458	890
12	Атропін	Навчальна	289,38	3	1	1,802	229709	7230	959
13	Азатіоприн	Навчальна	277,26	7	1	0,720	133848	4562	676
14	Баклофен	Навчальна	213,66	2	2	1,343	61670	2176	318

15	Бендрофлуметіазид	Навчальна	421,41	8	3	1,787	678665	11903	1752
16	Бетаметазон	Навчальна	392,47	6	3	1,143	573894	11755	1670
17	Бромокриптин	Навчальна	654,61	7	3	3,464	3108999	42905	5950
18	Буметанід	Навчальна	364,42	5	3	2,594	619100	10536	1446
19	Каптоприл	Навчальна	217,28	2	2	0,581	57629	2036	297
20	Карбамазепін	Навчальна	236,27	1	1	2,321	93062	4074	521
21	Цефалексин	Навчальна	347,39	4	3	0,008	427499	9735	1383
22	Хлортіазид	Навчальна	295,71	5	2	-0,506	93124	3142	460
23	Хлорталідон	Навчальна	338,76	4	3	1,318	249216	6678	952
24	Циметидин	Навчальна	252,34	6	3	1,140	190251	4813	664
25	Ципрофлоксацин	Навчальна	331,35	6	2	2,310	316562	8681	1234
26	Кладрибін	Навчальна	285,69	8	3	-0,065	132604	4399	671
27	Кліндаміцин	Навчальна	424,98	6	4	1,611	919724	12601	1843
28	Клоназепам	Навчальна	315,71	3	1	2,675	250800	6593	959
29	Клонідин	Навчальна	230,09	3	2	2,796	41807	2106	301
30	Клоксацилін	Навчальна	434,89	4	2	2,262	814301	15164	2212
31	Циклофосфамід	Навчальна	261,08	1	1	1,305	58317	1960	301
32	Цитарабін	Навчальна	243,22	6	4	-2,977	100180	3266	496

33	Дексаметазон	Навчальна	392,47	6	3	1,143	573894	11755	1670
34	Диданозин	Навчальна	236,23	6	2	-0,103	80271	3497	502
35	Дизопірамід	Навчальна	339,48	3	1	3,496	588748	10646	1372
36	Доксапрам	Навчальна	378,52	3	0	3,598	632487	14095	1843
37	Доксепін	Навчальна	279,38	2	0	3,983	210631	6973	882
38	Доксорубіцин	Навчальна	543,53	12	6	-0,381	2430223	29879	4392
39	Доксициклін	Навчальна	444,44	9	6	-1,055	1037402	16076	2336
40	Еналаприл	Навчальна	376,45	4	2	2,046	989295	14260	1982
41	Етинілестрадіол	Навчальна	296,41	2	2	4,192	204957	7264	941
42	Етопозид	Навчальна	588,56	12	3	1,102	2725244	36722	5464
43	Етідронат	Навчальна	206,03	3	5	-2,297	29666	814	136
44	Фамотидин	Навчальна	337,44	7	4	-1,791	414469	7046	1046
45	Флуконазол	Навчальна	306,28	9	1	0,290	261313	6772	1000
46	Фуросемід	Навчальна	330,74	5	3	1,220	286199	6372	939
47	Габапентин	Навчальна	171,24	2	2	1,327	27510	1432	193
48	Гемфіброзил	Навчальна	250,34	2	1	3,305	226682	5262	705
49	Гідралазин	Навчальна	160,18	4	2	0,671	18782	1417	182
50	Гідрохлоротіазид	Навчальна	297,73	5	3	-0,429	93124	3142	460



51	Ібупрофен	Навчальна	206,29	1	1	3,646	89861	3076	404
52	Ідарубіцин	Навчальна	497,50	10	5	0,320	1684951	24694	3545
53	Індапамід	Навчальна	365,83	4	2	2,188	416392	9719	1345
54	Ірбесартан	Навчальна	428,54	6	1	4,730	1188964	23576	3125
55	Ізосорбїду динїтрат	Навчальна	236,14	4	0	0,130	80397	2542	449
56	Ізосорбїду мононїтрат	Навчальна	191,14	4	1	-0,840	28699	1461	239
57	Кетамін	Навчальна	237,73	2	1	2,875	69755	2913	388
58	Кетопрофен	Навчальна	254,29	2	1	3,224	181720	5402	724
59	Лабеталол	Навчальна	328,41	4	4	2,524	636957	11893	1607
60	Лансопразол	Навчальна	369,36	8	1	1,493	545169	10846	1631
61	Леводопа	Навчальна	197,19	4	4	-1,757	62340	2160	321
62	Лїзиноприл	Навчальна	405,50	5	4	1,141	1355521	16912	2362
63	Лоразепам	Навчальна	321,16	3	2	3,479	196116	5637	819
64	Лозартан	Навчальна	422,92	7	2	4,846	1045260	19332	2665
65	Меперидин	Навчальна	247,34	2	0	2,609	128001	4369	568
66	Меркаптопурин	Навчальна	152,18	4	2	0,156	7868	798	108
67	Метформін	Навчальна	129,17	5	4	0,607	13687	716	96
68	Метотрексат	Навчальна	454,45	10	5	1,405	2190888	26126	3832

69	Метилдопа	Навчальна	211,22	4	4	-1,632	83432	2552	374
70	Метилпреднізолон	Навчальна	374,48	5	3	1,507	498115	11202	1553
71	Метоклопрамід	Навчальна	299,80	4	2	1,852	357018	6445	902
72	Метопролол	Навчальна	267,37	4	2	1,643	322987	6707	906
73	Метронідазол	Навчальна	171,16	4	1	-0,123	27473	1288	193
74	Мексилетин	Навчальна	179,26	2	1	2,081	42838	2034	256
75	Мізопростол	Навчальна	382,54	4	2	3,449	1677329	16886	2331
76	Налоксон	Навчальна	327,38	5	2	0,612	232022	7720	1047
77	Налтрексон	Навчальна	341,41	5	2	0,705	252019	8803	1189
78	Ніфедипін	Навчальна	346,34	4	1	1,953	549472	8792	1280
79	Німодипін	Навчальна	418,45	5	1	2,667	1401983	15835	2290
80	Нітразепам	Навчальна	281,27	3	1	2,053	204937	6142	857
81	Нізатидин	Навчальна	331,45	5	2	1,156	535958	8353	1227
82	Норфлуксацин	Навчальна	319,34	6	2	2,211	317490	7795	1116
83	Нортриптилін	Навчальна	263,38	1	1	4,683	165063	6143	759
84	Омепразол	Навчальна	345,42	6	1	0,194	438619	10452	1419
85	Ондансетрон	Навчальна	293,37	4	0	2,508	217302	7706	997
86	Орципреналін	Навчальна	211,26	4	4	0,999	87750	2864	394

87	Оксспренолол	Навчальна	265,35	4	2	2,036	297610	6263	834
88	Оксикодон	Навчальна	315,37	5	1	0,306	186078	6673	907
89	Памідронат	Навчальна	235,07	4	6	-3,141	67926	1345	220
90	Пантопразол	Навчальна	383,37	9	1	0,117	650686	11929	1806
91	Петидин	Навчальна	247,34	2	0	2,609	128001	4369	568
92	Фенобарбітон	Навчальна	232,24	3	2	1,327	92531	3322	458
93	Фенітоїн	Навчальна	252,27	2	2	2,252	121875	4679	617
94	Піндолол	Навчальна	248,33	4	3	1,992	155129	5206	687
95	Правастатин	Навчальна	424,53	5	4	2,133	1592942	18188	2600
96	Празозин	Навчальна	383,41	8	1	1,312	754065	15647	2195
97	Примідон	Навчальна	218,26	2	2	1,526	69194	2898	385
98	Прокаїнамід	Навчальна	235,33	3	2	1,328	176834	4715	618
99	Прохлорперазин	Навчальна	373,94	3	0	4,544	427920	11513	1544
100	Прометазин	Навчальна	284,42	2	0	4,371	162136	5850	744
101	Пропілтіоурацил	Навчальна	170,23	1	1	0,477	18909	1161	158
102	Квінаприл	Навчальна	438,52	4	2	3,139	1650793	22541	3064
103	Хінін	Навчальна	324,42	4	1	2,584	330149	9833	1286
104	Ранітидин	Навчальна	314,40	5	2	1,301	535958	8479	1227

105	Рибавірин	Навчальна	244,21	8	4	-2,875	103957	3311	515
106	Саліцилова кислота	Навчальна	138,12	2	2	2,208	11308	822	114
107	Селегілін	Навчальна	187,29	1	0	2,687	65396	2766	337
108	Симвастатин	Навчальна	418,57	3	1	4,649	1159529	17942	2440
109	Соталол	Навчальна	272,36	4	3	1,079	226065	5254	706
110	Спіронолактон	Навчальна	416,58	3	0	3,390	601303	13884	1905
111	Сульфометоксазол	Навчальна	253,28	5	2	0,690	108079	3926	535
112	Сульфізоксазол	Навчальна	267,30	5	2	0,913	138704	4543	614
113	Суматриптан	Навчальна	295,40	4	2	0,908	237792	6516	859
114	Такрин	Навчальна	198,27	2	1	2,205	41269	2657	326
115	Темазепам	Навчальна	300,74	3	1	2,997	194836	5895	814
116	Тербуталін	Навчальна	225,29	4	4	1,414	118746	3420	468
117	Тестостерон	Навчальна	288,43	2	1	3,535	160228	6172	802
118	Тетрациклін	Навчальна	444,44	9	6	-0,762	1043893	16144	2350
119	Теофілін	Навчальна	180,17	4	1	-0,746	25429	1509	211
120	Тимолол	Навчальна	316,42	7	2	1,309	324259	7477	1063
121	Толбутамід	Навчальна	270,35	3	2	2,141	219163	5180	685
122	Триамтерен	Навчальна	253,27	7	3	1,693	132417	5012	670

123	Тримеперидин	Навчальна	294,44	2	0	4,265	255824	7852	979
124	Трандолаприл	Навчальна	402,49	4	3	2,610	1041851	17000	2340
125	Тропісетрон	Навчальна	284,36	3	1	3,047	188872	7235	944
126	Вальпроєва кислота	Навчальна	144,21	1	1	2,661	23285	1000	131
127	Верапаміл	Навчальна	454,61	6	0	4,665	2740315	27284	3698
128	Варфарин	Навчальна	308,33	3	1	2,985	310959	8146	1092
129	Зальцитабін	Навчальна	211,22	4	2	-1,394	58431	2583	369
130	Зидовудин	Навчальна	267,25	6	2	-0,884	175676	4676	701
131	Амантадин	Тестова	151,25	1	1	2,236	8844	1067	124
132	Аміодарон	Тестова	645,32	4	0	7,861	1463147	21156	2897
133	Амітриптилін	Тестова	277,41	1	0	4,932	210631	7133	882
134	Хлорамфенікол	Тестова	323,13	4	3	0,669	348207	5269	880
135	Хлордіазепоксид	Тестова	299,76	3	1	2,729	198023	6091	828
136	Діазепам	Тестова	284,74	2	0	2,921	158292	5393	726
137	Доласетрон	Тестова	324,38	4	1	2,217	303937	10213	1368
138	Етамбутол	Тестова	204,31	4	4	-0,360	138073	2798	383
139	Ламівудин	Тестова	229,25	4	2	-1,451	58431	2479	369
140	Ламотриджин	Тестова	256,09	5	2	3,876	74968	2813	417

141	Флунітразепам	Валідаційна	313,29	4	0	2,354	302738	7365	1063
142	Лідокаїн	Валідаційна	234,34	2	1	2,110	159216	4344	556
143	Раміприл	Валідаційна	416,52	4	2	2,797	1210619	18583	2546
144	Морфін	Валідаційна	285,34	4	2	0,775	123644	5332	712
145	Міноциклін	Валідаційна	457,48	9	5	0,110	1258222	18638	2672

\* 1 – молекулярна маса, 2 – логарифмічний коефіцієнт розподілу, 3 – кількість акцепторів водневого зв'язку, 4 – кількість донорів водневого зв'язку, 5 – індекс Балабана, 6 – молекулярний топологічний індекс, 7 –індекс Вінера.

Встановлення правильної кількості прихованих нейронів у нейронній мережі є критично важливим для досягнення оптимальної продуктивності моделі. Це рішення впливає на здатність моделі навчатися, узагальнювати знання та ефективно працювати з новими даними [30, 31].

Важливо правильно визначити кількість прихованих нейронів:

– уникнення перенавчання (overfitting) – якщо прихованих нейронів занадто багато, модель може стати надто складною. Вона починає «запам'ятовувати» навчальні дані замість узагальнення закономірностей, що призводить до поганої роботи на тестових даних.

– уникнення недонавчання (underfitting) – якщо прихованих нейронів недостатньо, модель може бути занадто простою, щоб навчитися складним залежностям у даних, що також знижує її продуктивність.

У цій роботі оптимальну кількість прихованих нейронів знайдено експериментальним шляхом. Штучна нейронна мережа прямого поширення сигналу була реалізована при різній кількості нейронів у прихованому шарі (від 7 до 18). Оптимальна кількість прихованих нейронів забезпечує належне навчання штучних нейронних мереж. Для кожної кількості прихованих нейронів оцінювали коефіцієнт детермінації між експериментальними та прогнозованими значеннями біодоступності для 130 лікарських речовин навчальної вибірки.

Залежність ефективності навчання мережі від кількості прихованих нейронів представлено на рисунку 1.2.



**Рисунок 1.2.** Залежність коефіцієнта детермінації між експериментальними та прогнозованими значеннями біодоступності для навчальної вибірки від кількості прихованих нейронів для нейронної мережі прямого поширення сигналу

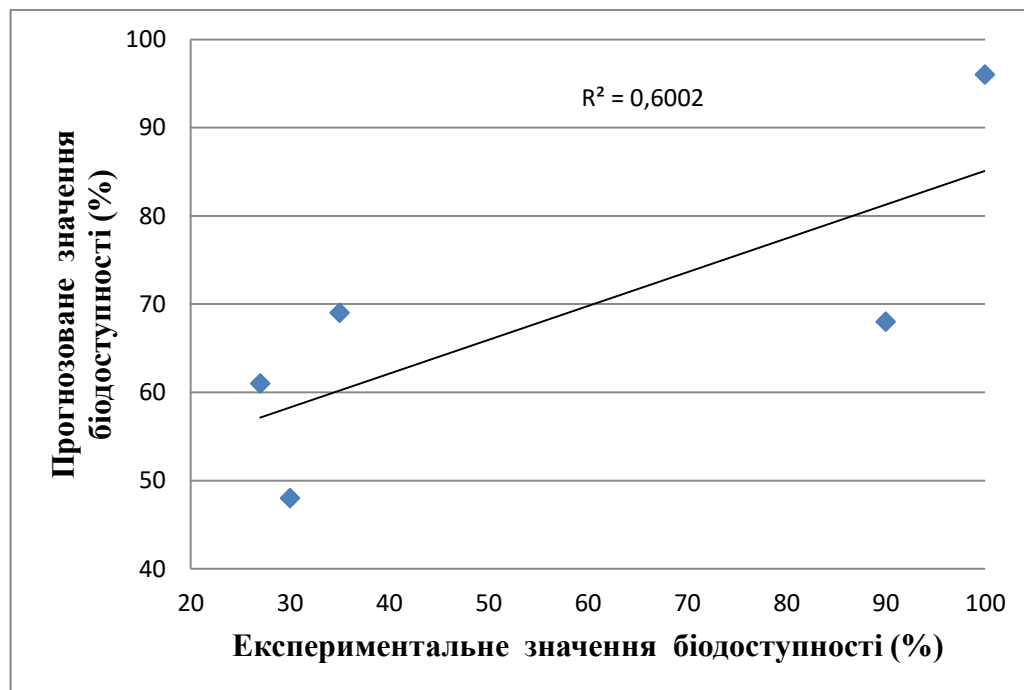
Оптимальна кількість прихованих нейронів для нейронної мережі прямого поширення сигналу становить 16, оскільки при цій кількості прихованих нейронів спостерігається максимальне значення коефіцієнта детермінації ( $R = 0,3160$ ), що свідчить про найкраще «засвоєння» інформації нейронною мережею.

#### **2.4 Результати прогнозування біодоступності лікарських речовин із застосуванням нейронної мережі прямого поширення сигналу**

Навчена нейронна мережа прямого поширення сигналу із 16 прихованими нейронами використовувалася для прогнозування значень біодоступності для лікарських речовин валідаційної вибірки та тестової вибірки.



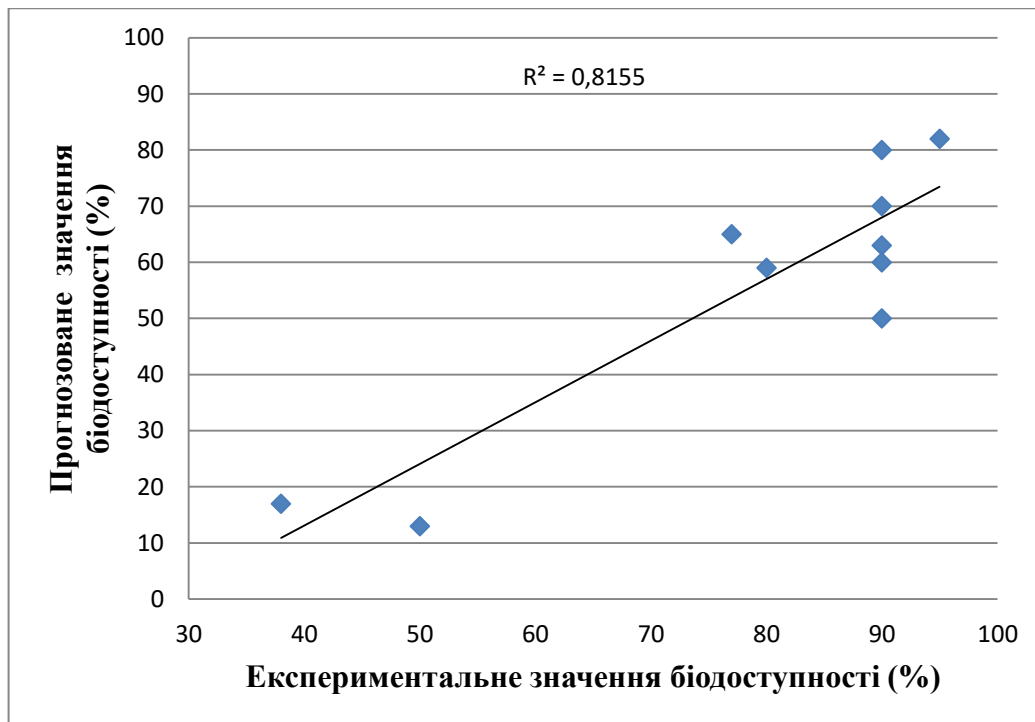
Залежність прогнозованих значень біодоступності від експериментальних значень біодоступності для валідаційної вибірки представлена на рисунку 2.2.



**Рисунок 2.2.** Залежність прогнозованих значень біодоступності від експериментальних значень біодоступності для лікарських речовин валідаційної вибірки

Хоча коефіцієнт детермінації 0,6002 не є достатньо високим, але у випадку аналізу складних залежностей як прогнозування біодоступності лікарських речовин на основі їх хімічної структури він свідчить про адекватну якість моделі.

Залежність прогнозованих значень біодоступності від експериментальних значень біодоступності для тестової вибірки представлена на рисунку 3.2.



**Рисунок 3.2.** Залежність прогнозованих значень біодоступності від експериментальних значень біодоступності для лікарських речовин тестової вибірки

Високе значення коефіцієнта детермінації ( $R = 0,8155$ ) свідчить про можливість її застосування для прогнозування біодоступності лікарських речовин.

Перспективи подальших досліджень полягають у розширенні навчальної вибірки та у включенні нових дескрипторів для оцінки їх інформативності. Ці кроки мають забезпечити більш точніші навчання нейронної мережі та прогнозування біодоступності нових потенційно перспективних лікарських молекул.

## ВИСНОВКИ

1. Встановлено, що з 16 досліджених молекулярних дескрипторів лише 7 є інформативними та достатніми для прогнозу біодоступності лікарських речовин для перорального застосування: молекулярна маса, логарифмічний коефіцієнт розподілу, кількість акцепторів водневого зв'язку, кількість донорів водневого зв'язку та топологічні індекси (індекс Балабана, молекулярний топологічний індекс, індекс Вінера).

2. Показано, що нейронна мережа прямого поширення сигналу із 16 прихованими нейронами у своїй структурі є надійним інструментом для прогнозування біодоступності лікарських речовин на основі їх хімічної структури.

3. Запропонований алгоритм поєднання набору із 7 молекулярних дескрипторів та нейронної мережі для прогнозування біодоступності лікарських речовин може бути корисним на ранніх етапах відбору перспективних молекул-кандидатів.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Toutain, P. L., & Bousquet-mélou, A. (2004). Bioavailability and its assessment. *Journal of veterinary pharmacology and therapeutics*, 27(6), 455-466.
2. Allam, A. N., El Gamal, S., & Naggar, V. (2011). Bioavailability: A pharmaceutical review. *Int J Novel Drug Deliv Tech*, 1(1), 77-93.
3. Abo-EL-Sooud, K. (2020). Absolute and relative bioavailability. In *Drug Discovery and Evaluation: Methods in Clinical Pharmacology* (pp. 879-885). Cham: Springer International Publishing.
4. Stielow, M., Witczyńska, A., Kubryń, N., Fijałkowski, Ł., Nowaczyk, J., & Nowaczyk, A. (2023). The bioavailability of drugs — the current state of knowledge. *Molecules*, 28(24), 8038.
5. Khadka, P., Ro, J., Kim, H., Kim, I., Kim, J. T., Kim, H., ... & Lee, J. (2014). Pharmaceutical particle technologies: An approach to improve drug solubility, dissolution and bioavailability. *Asian journal of pharmaceutical sciences*, 9(6), 304-316.
6. Gomez-Orellana, I. (2005). Strategies to improve oral drug bioavailability. *Expert opinion on drug delivery*, 2(3), 419-433.
7. Naidu, R., Semple, K. T., Megharaj, M., Juhasz, A. L., Bolan, N. S., Gupta, S. K., ... & Schulin, R. (2008). Bioavailability: definition, assessment and implications for risk assessment. *Developments in soil science*, 32, 39-51.
8. Józsa, L., Nemes, D., Pető, Á., Kósa, D., Révész, R., Bácskay, I., ... & Vasvári, G. (2023). Recent options and techniques to assess improved bioavailability: In vitro and ex vivo methods. *Pharmaceutics*, 15(4), 1146.
9. Becker, C., Boettcher, M., Muenster, U., Loewen, S., Lobmeyer, M., & Mueck, W. (2022). Results from in vitro and in vivo studies evaluating the bioavailability, effects of food, and administration as crushed tablet suspension on vericiguat pharmacokinetics. *AAPS Open*, 8(1), 16.

10. Fahmy, S., & Abu-Gharbieh, E. (2014). In vitro dissolution and in vivo bioavailability of six brands of ciprofloxacin tablets administered in rabbits and their pharmacokinetic modeling. *BioMed research international*, 2014(1), 590848.
11. Alqahtani, S. (2023). Improving on in-silico prediction of oral drug bioavailability. *Expert Opinion on Drug Metabolism & Toxicology*, 19(10), 665-670.
12. Wei, M., Zhang, X., Pan, X., Wang, B., Ji, C., Qi, Y., & Zhang, J. Z. (2022). HobPre: Accurate prediction of human oral bioavailability for small molecules. *Journal of Cheminformatics*, 14(1), 1.
13. Yu, L. X., Ellison, C. D., & Hussain, A. S. (2004). Predicting human oral bioavailability using in silico models. In *Applications of pharmacokinetic principles in drug development* (pp. 53-74). Boston, MA: Springer US.
14. Ye, Z., Yang, Y., Li, X., Cao, D., & Ouyang, D. (2018). An integrated transfer learning and multitask learning approach for pharmacokinetic parameter prediction. *Molecular pharmaceutics*, 16(2), 533-541.
15. Han, S. H., Kim, K. W., Kim, S., & Youn, Y. C. (2018). Artificial neural network: understanding the basic concepts without mathematics. *Dementia and neurocognitive disorders*, 17(3), 83-89.
16. Walczak, S. (2018). Artificial neural networks. In *Encyclopedia of Information Science and Technology, Fourth Edition* (pp. 120-131). IGI Global Scientific Publishing.
17. Kudus, S. A., Bunnori, N. M., Basri, S. R., Shahiron, S., Jamil, M. N. M., & Noorsuhada, M. N. (2013). An overview current application of artificial neural network in concrete. *Advanced Materials Research*, 626, 372-375.
18. Wu, Y. C., & Feng, J. W. (2018). Development and application of artificial neural network. *Wireless Personal Communications*, 102, 1645-1656.
19. The Prime Chemistry Portal. URL: <https://chemistrydocs.com/perkinelmer-chemoffice-2020-version-20-0/>.
20. Matlab for artificial intelligence. URL: <https://www.mathworks.com/products/matlab.html>.

21. Sazlı, M. H. (2006). A brief review of feed-forward neural networks. *Communications Faculty of Sciences University of Ankara Series A2-A3 Physical Sciences and Engineering*, 50(01).
22. Mumali, F. (2022). Artificial neural network-based decision support systems in manufacturing processes: A systematic literature review. *Computers & Industrial Engineering*, 165, 107964.
23. Kelley, K., & Bolin, J. H. (2013). Multiple regression. In *Handbook of quantitative methods for educational research* (pp. 69-101). Brill.
24. Turner, J. V., Glass, B. D., & Agatonovic-Kustrin, S. (2003). Prediction of drug bioavailability based on molecular structure. *Analytica Chimica Acta*, 485(1), 89-102.
25. Yang, Q., Fan, L., Hao, E., Hou, X., Deng, J., Xia, Z., & Du, Z. (2024). Construction of An Oral Bioavailability Prediction Model Based on Machine Learning for Evaluating Molecular Modifications. *Journal of Pharmaceutical Sciences*, 113(5), 1155-1167.
26. Chagas, C. M., Moss, S., & Alisaraie, L. (2018). Drug metabolites and their effects on the development of adverse reactions: Revisiting Lipinski's Rule of Five. *International journal of pharmaceutics*, 549(1-2), 133-149.
27. Ivanović, V., Rančić, M., Arsić, B., & Pavlović, A. (2020). Lipinski's rule of five, famous extensions and famous exceptions. *Popular Scientific Article*, 3(1), 171-177.
28. Burzykowski, T., Geubelmans, M., Rousseau, A. J., & Valkenburg, D. (2023). Validation of machine learning algorithms. *American Journal of Orthodontics and Dentofacial Orthopedics*, 164(2), 295-297.
29. Pushkarova, Y., Zaitseva, G., & Kaliuzhenko, A. (2023, September). Classification of Residual Solvents by Risk Assessment Using Chemometric Methods. In *2023 13th International Conference on Advanced Computer Information Technologies (ACIT)* (pp. 562-565). IEEE.

30. Rachmatullah, M. I. C., Santoso, J., & Surendro, K. (2021). Determining the number of hidden layer and hidden neuron of neural network for wind speed prediction. *PeerJ Computer Science*, 7, e724.

31. Pushkarova, Y., Zaitseva, G., & Al Saker, M. (2022, September). Prediction of Toxicity of Phenols Using Artificial Neural Networks. In *2022 12th International Conference on Advanced Computer Information Technologies (ACIT)* (pp. 493-496). IEEE.

## ДОДАТОК



# СЕРТИФІКАТ

№18/2025.S2.07

Цим засвідчується, що

**Тимченко І.І.**

з постерною доповіддю

**ПРОГНОЗУВАННЯ БІОДОСТУПНОСТІ ЛІКАРСЬКИХ РЕЧОВИН НА ОСНОВІ ЇХ ХІМІЧНОЇ СТРУКТУРИ**

брала участь у Науково-практичній конференції з міжнародною участю

**«PLANTA+. НАУКА, ПРАКТИКА ТА ОСВІТА»**

присвяченій пам'яті доктора хімічних наук, професорки Ніни Павлівни Максютіної (до 100-річчя від дня народження) в онлайн-форматі

28–29 січня 2025 р., м. Київ, Україна



# PLANTA+

НАУКА, ПРАКТИКА ТА ОСВІТА

Конференція організована у співпраці з науковцями Інституту фармакології та токсикології Національного медичного університету імені О.О. Богомольця (Київ, Україна) (№ 021/11-19-021 від 13 листопада 2024 р.)

Ректор Національного медичного університету імені О. О. Богомольця, д. м. н., професор



Юрій КУЧИН

Завідувачка кафедри фармакогнозії та ботаніки, д. б. н., професор

Валентина МІНАРЧЕНКО



# СЕРТИФІКАТ

№298/2025

Цим засвідчується, що

**Тимченко І. І.**

брав(-ла) участь у Науково-практичній конференції з міжнародною участю

**«PLANTA+. НАУКА, ПРАКТИКА ТА ОСВІТА»,**

присвяченій пам'яті доктора хімічних наук, професорки Ніни Павлівни Максютіної (до 100-річчя від дня народження)

Тривалістю 6 годин (0,2 кредита ЕКТС)

28–29 січня 2025 р., м. Київ, Україна



# PLANTA+

НАУКА, ПРАКТИКА ТА ОСВІТА

Конференція організована у співпраці з науковцями Інституту фармакології та токсикології Національного медичного університету імені О.О. Богомольця (Київ, Україна) (№ 021/11-19-021 від 13 листопада 2024 р.)

Ректор Національного медичного університету імені О. О. Богомольця, д. м. н., професор



Юрій КУЧИН

Завідувачка кафедри фармакогнозії та ботаніки, д. б. н., професор

Валентина МІНАРЧЕНКО



## SUMMARY

**Inara Tymchenko**

### **PREDICTION OF ORAL DRUG BIOAVAILABILITY BASED ON CHEMICAL STRUCTURE**

**Department of Analytical, Physical and Colloid Chemistry**

**Scientific supervisor:** Candidate of Chemical Sciences, Associate Professor Yaroslava Pushkarova

**Keywords:** chemometrics, drug design, molecular descriptor, multivariate regression, neural network, pharmacy.

**Introduction.** Prediction of bioavailability is a crucial aspect of drug development and formulation design as it directly influences the therapeutic efficacy of a drug. Bioavailability refers to the fraction of an administered dose of a drug that reaches the systemic circulation in a therapeutically active form. Accurate prediction of bioavailability is essential in determining whether a drug will achieve the desired therapeutic effects when administered via the oral route, which is the most common method of drug administration. **Aim.** To investigate the possibility of predicting the bioavailability of oral drug compounds based on molecular descriptors by means of feedforward neural network.

**Materials and methods.** Software ChemOffice 2020 was used for calculation molecular descriptors from the drug structures. Software Matlab R2024b was used for building the multivariate regression between molecular descriptors and bioavailability as well as feedforward neural network. For each of the 145 drug molecules 16 molecular descriptors were calculated. Of the 145 compounds, 10 were randomly selected for use as a test subset, 5 compounds – as a validation subset. The remaining 130 compounds were used to train feedforward neural network.

**Results and discussion.** It is established that there are 7 the most informativeness molecular descriptors for oral drug bioavailability prediction: molecular weight, hydrogen bond acceptors, hydrogen bond donors, partition coefficient, Balaban index, molecular topological index, Wiener index (multiple

coefficient of determination is equal to 0,3701). Optimal number of hidden neurons for effective realization feedforward neural network was found by experimental way and its equal to 16. Trained feedforward neural network with 16 hidden neurons was used for predicting bioavailability values for test and validation subsets.

**Conclusions.** Feedforward neural network is the effective tool for prediction of oral drug bioavailability based on chemical structure. This is evidenced by high values of determination coefficient between predicted bioavailability values and experimental bioavailability values for test and validation subsets (0,7084 and 0,8432, correspondingly). Obtained results can be useful at the stage of experiment planning or drug design.