

Т. Булана, Б. Молодець

**АНАЛІЗ ІСНУЮЧИХ ВАРІАНТІВ КЛАСИФІКАЦІЇ ХВОРИХ
НА СЕРЦЕВО-СУДИННИМИ ЗАХВОРЮВАННЯМИ
ЗА ДОПОМОГОЮ НЕЙРОННИМИ МЕРЕЖАМИ**

Анотація. Робота присвячена аналізу інформаційних технологій хронобіологічного моніторингу кардіосистем, розробці систему підтримки прийняття рішень для лікаря-дослідника на базі методів класифікації з використанням нейронних мереж таких як імовірнісна неронна мережа PNN (Probabilistic Neural Networks), багатопшаровий перцептрон MLP NN (Multi-Layer Perceptron), каскадно-кореляційна мережа CasCor (Cascade Correlation). У результаті отримано наступне: найкращим класифікатором є нейромережа каскадної кореляції з 85-88% точністю класифікації. Найгіршим класифікатором стала ймовірнісна нейронна мережа, оскільки точність цього алгоритму залежить від розміру набору даних.

Ключові слова: хронобіологія, інформаційна технологія моніторингу, нейронні мережі, класифікація.

В наш час науково-технічний прогрес увійшов майже у всі сфери нашого життя. Насичення всіх галузей сучасного життя комп'ютерними технологіями не залишило осторонь і медицину. Зі зростанням темпу життя від лікаря часто вимагається оперативності в постановці діагнозу та виборі оптимального і сучасного виду лікування. Зрозуміло, що для виконання таких непростих завдань необхідні відповідні інструменти. Медична культура й охорона здоров'я нині на порозі технічної революції. Інформаційні технології стрімко вторгаються в усі сфери медицини й організації охорони здоров'я. Сьогодні практично жоден етап діагностики не обходиться без комп'ютерних технологій. Поряд з цим, інтелектуальні інформаційні системи доволі обмежено застосовують у практичній медицині. Застосування інтелектуальних систем в медицині, без сумніву, сприяє прогресивному розвитку інформаційного потенціалу,

який є універсальним засобом вирішення широкого кола завдань на різних етапах лікування пацієнта [1].

В ХХ столітті вивчення часових особливостей біологічних структур виділилось у два нові напрями науки – хрономедицину та хронобіологію. В 1937 році у шведському курортному місті Роннебі вперше зібрались представники різноманітних теоретичних та практичних дисциплін та заснували міжнародну спільноту по вивченню біологічних ритмів. Тим самим вони започаткували розвиток науки та розширенню числа розуміючих важливість часового фактору для життя. Одне з ключових положень хронобіології: часовий фактор у вигляді у вигляді ритмів характерний для всіх процесів життєдіяльності. Таким чином, можна сказати що хронобіологія – це дослідження ритмів[2].

На сьогодні через досить примітивну будівлю програмного забезпечення для приладів-реєстраторів не дозволяє проводити детальний комплексний аналіз оцінок, показників і характеристик кардіосистем, який необхідний для виявлення прихованих закономірностей; в основному діагностика ґрунтується на досвіді і знанні лікаря; неясно, який обсяг показників добового моніторингу артеріального тиску необхідно використовувати для оцінки ефективності проведеного лікування і не існує системи, яка б дозволяла оцінити якість проведеної лікувальної терапії. Таким чином, прийняття рішення виглядає наступним чином (рис 1.).

Для вирішення цієї проблеми було вирішено використати провести аналіз деяких нижче перелічених мереж: метод каскадної кореляції (CasCor), багатошаровий перцептрон (MLP), та ймовірнісна нейронної мережі (PNN). В якості методу відбору інформативних ознак було обрано множинну логістичну регресію. Перевагою запропонованого методу відбору інформативних ознак є скорочення обсягу інформації: вплив кожного фактора шляхом оцінювання коефіцієнтів регресійного рівняння. Основним недоліком логістичної регресії є умова, що змінна відгуку повинна бути бінарною.

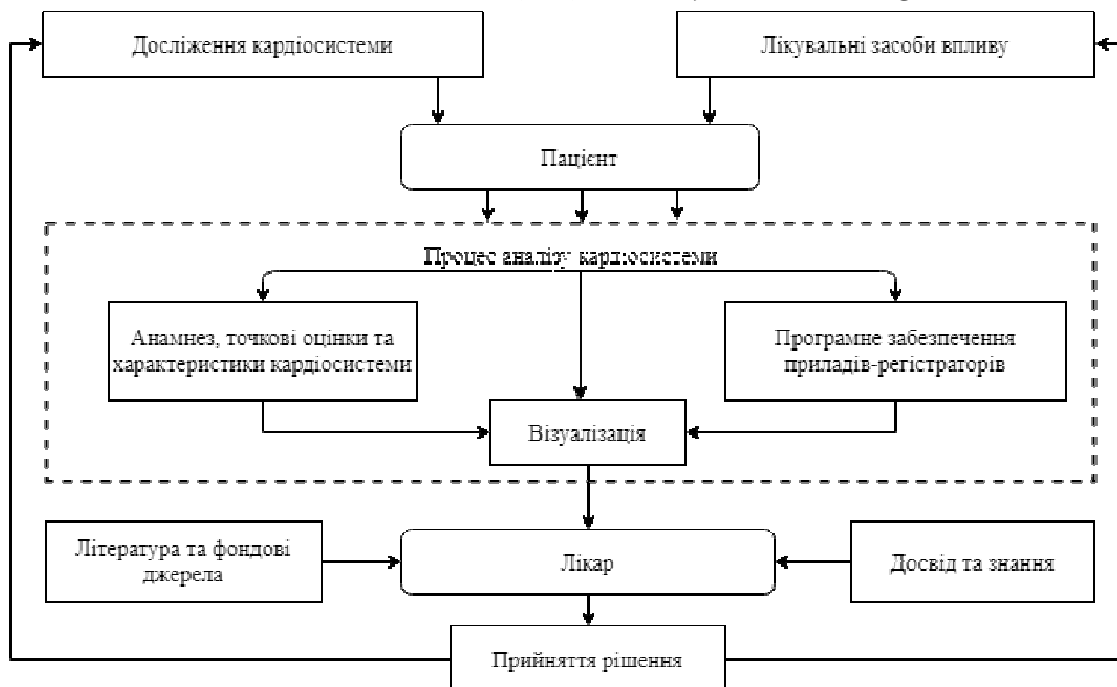


Рисунок 1 - Схема проведення діагностики стану кардіосистеми пацієнта

Багатошаровий перцептрон найбільш вдало реалізує ідею керованого навчання, в якому розглядається відображення вхідного вектора в цільовий вихідний вектор. Але для вдалого управління навчанням необхідно мати дані, які мають відому класифікацію, і обсяг, достатній для проведення навчання і тестування НМ.

В свою чергу Імовірнісні нейронні мережі (PNN, Probabilistic Neural Networks), також відносяться до НС з прямою передачею інформації і по архітектурі є «найближчими родичами» радіально-базисних і узагальнених регресійних мереж Дані мережі призначені для вирішення завдань байесівської класифікації (розпізнавання образів на основі байесівського підходу), в основі якої лежить формула Байеса[3]. Особливістю цієї мережі є фіксована структура, що будується один раз по набору навчальних пар. В процесі функціонування ваги синапсів не змінюються. До нейронних мереж з прямою передачею інформації відносяться і так звані каскадно-кореляційні мережі, основна особливість яких полягає в можливості додавання нових вузлів в процесі навчання.

Алгоритм каскадної кореляції комбінує дві ключові ідеї: перша - це каскадна архітектура, в якій приховані нейрони додаються до мережі один раз і не змінюються після того, як вони були додані[4]; друга - навчальний алгоритм, який створює і налаштовує нові приховані нейрони (рис. 2). Для кожного нового прихованого нейрона, необхідно максимізувати величину кореляції між виходом нового нейрона і сигналом похибки, який необхідно знизити.

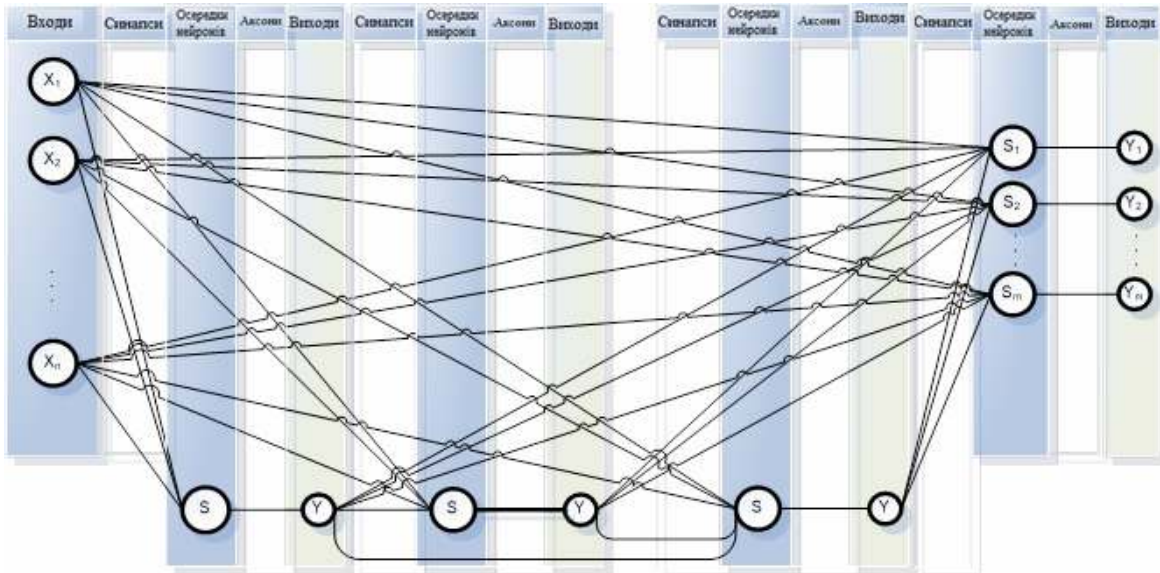


Рисунок 2 - Загальна архітектура нейронної мережі каскадної кореляції

Вхідні данні, що надходять до нашої програми можна розділити на первинні неопрацьовані данні та данні опрацьовані лікарем-дослідником. Відмінною характеристикою даних є їх частота отримання. Апарат вимірювання артеріального тиску закріплюється на пацієнта на добу. Частота забору – чотири заміри за одну годину, тобто 96 на добу. Переваги такого хронобіологічного моніторингу та збору даних є відсутність ефекту «білого халату» - данні збираються не у стаціонарному режимі, тобто пацієнт не повинен знаходитись на території медичного закладу.

Загальна кількість оцінок, наданих лікарями-дослідниками для проведення хронобіологічного аналізу, становить 157 серед яких такі кількісні показники: САТ (систоличний артеріальний тиск) мінімум за

добу, САТ мінімум за добу, САТ мінімум вночі, САТ індекс площі гіпотонії за добу та інше (рис. 3). У пацієнта також є такі якісні параметри як вага, зріст, дата звернення до лікарні, стаж хвороби, діагноз, менопауза. Серед якісних показників є вживання алкоголю, куріння, спадкова схильність до інфаркту міокарда. Нормування навчальних даних проводилось в інтервалі [-1;1]. А для класифікації використовувалось кількість вхідних параметрів в розмірі восьми оцінок. Тип вхідних оцінок – числовий (ДАТ середнє за добу, ПАТ середнє за добу, САТ середнє за добу та інші). Навчальна вибірка склала 80% від загальної кількості хворих (278 навчальних пар), а тестова вибірка – 20% (65 навчальних пар). Представлена вибірка даних є високорепрезентативною, враховуючи кількість пацієнтів: всього внесено даних по 343 хворих з серцево-судинними ускладненнями (1 клас – 142, 2 клас – 207).

Характеристики	
САТ середнє за добу	157 параметрів
САТ середнє за ніч	
САТ середнє за день	
САТ мінімум за ніч	
...	
Пульс МАХ (макс. ранковий підйом)	
Пульс середній за день	

Рисунок 3 - Таблиця ознак у вхідній вибірці

Але на даних, що отримуються в результаті цього моніторингування, існують похибки трьох типів: похибка виміру приладу, похибка взаємодії пацієнта з приладом, похибка обчислювального процесу.

Одна з головних причин виникнення похибки першого типу є несправність приладу, бо у кожного з них є певні гарантійні терміни роботи, які не завжди збігаються з реальністю.

Другий тип похибку обумовлений процедурою моніторингу та машино-людина відносинами. Щоб розуміти цю похибку треба ознайомитись з принципами роботи приладу моніторингу артеріального тиску: раз в N хвилин (де N – довільна величини, що залежить від марки приладу) без попереднього оповіщення пацієнта, починається процедура заміру – починається стискуватись манжетка; якщо пацієнт в стані неспокою(не фіксує руку – що є необхідним для отримання точного результату), прилад продовжує накачувати манжетку штучно спотворюючи цей результат. Під час сну деякі пацієнти можуть знімати прилади або послаблювати манжетку, що призведе до спотворення даних. Також данні спотворює емоційний стан людини.

Третя похибка виникає у будь-якому виді мануального заповнення даних до БД лікарем-дослідником.

Спираючись на вище описані спостереження, було вирішено граничні значення АТ досліджувати окремо.

У результаті отримано наступне: найкращим класифікатором є нейромережа каскадної кореляції з 85-88% точністю класифікації. Найгіршим класифікатором стала ймовірнісна нейронна мережа, оскільки точність цього алгоритму залежить від розміру набору даних - якість буде кращою, якщо використати великий за обсягом датасет. Також потрібно зауважити, що віковий діапазон пацієнтів у вибірці 50-60 років. З вище описаних зауважень можна зробити висновки про необхідність накопичення даних для збільшення навчальної вибірки.

Схожі дослідження були проведені співробітниками Ліонський університет І імені Клода Бернара. Для класифікації пацієнтами потенційно схильними до інфаркту міокарда вони використовували дерево рішень, (CART) та багат шаровий перцептрон (MLP) та логістичної регресії (Logistic Regression). В якості даних для навчання та тестування були використані данні з бази даних INDANA (INdividual Data ANALysis of Antihypertensive interventions). Ця вибірка була випадковим чином розбита на навчальну (кількість записів – 10,296) та тестову (кількість записів – 5,148). Ознаки пацієнтів які приймали участь у навчанні: вік, стать, систолічний артеріальний тиск (CAT), діастолічний

артеріальний тиск (ДАТ), захворювання діабетом, наявність шкідливих звичок (куріння), гіпертрофія лівого шлуночка на ЕКГ, індекс маси тіла та рівень холестерину. У результаті були отримане наступне: найбільш інформативними ознаками виявились вік та стать пацієнта. Якість розпізнання «уражених» пацієнтів наступні: логістична регресія (LR) – 65.9%, дерево прийняття рішень (CART) – 69.1%, багатосаровий перцептрон (MLP) – 76% [5].

ЛІТЕРАТУРА / LITERATURE

1. Дзяк Г. В. Суточное мониторирование артериального давления / Г. В. Дзяк, Т. В. Колесник, Ю. Н. Погорецкий. – Днепропетровск, 2005. – 200 с.
2. Г. Хильдебрандт , Хронобиология и хрономедицина / Г. Хильдебрандт, М. Мозер, М.Лехофер. М.: Арнебия. 2006. – 144с.
3. Буланая Т.М. Диагностика сердечно-сосудистой системы на основе вероятностных нейронных сетей / Т. М. Булана // Математичне та програмне забезпечення інтелектуальних систем Мрзис-2007: міжн. наук.-практ. конф., 14–16 листопада 2007 р.: тези доп. – Дніпропетровськ, 2007.– С. 31.
4. Булана Т.М. Інформаційні технології моніторингу й оцінки ефективності діагностування стану технічних та кардіосистем: Дис, ... кандидата тех. наук: 05.13.06: – Захищене 27.06.2018; Затв. 23.10.2018. – Д., 2018. – 145 с.
5. Colombet I., Ruelland A., Chatellier G., Gueyffier F., Degoulet P., Jaulent M.C. Models to predict cardiovascular risk: Comparison of CART, multilayer perceptron and logistic regression. Proceedings AMIA Symposium , 2000. 156–160.

REFERANCE

1. Dziak G.V. Daily blood pressure monitoring / G.V. Dziak, T.V. Kolesnik, Y.N. Pogoretskii. – Dnipropetrovsk, 2005. – 200 p.
2. H. Khildebrandt , Chronobiology and Chronomedicine / H. Khildebrandt, M. Mozer, M. Lekhofer. M.: Arnebia. 2006. – 144p.
3. Bulanaya T.M. Diagnosis of the cardiovascular system based on probabilistic neural networks / T. M. Bulanaya // Мрзис-2007: international Research Practice Conf., 14 – 16 november 2007 p.: thesis. – Dnipropetrovsk, 2007. – p. 31.
4. Bulanaya T.M. Information technologies for monitoring and evaluating the efficiency of diagnosing the condition of technical and cardio systems: Dis, ...

Candidate of Technical Sciences: 05.13.06: – Defended 27.06.2018; Approved 23.10.2018. – D., 2018. – 145 p.

5. Colombet I., Ruelland A., Chatellier G., Gueyffier F., Degoulet P., Jaulent M.C. Models to predict cardiovascular risk: Comparison of CART, multilayer perceptron and logistic regression. Proceedings AMIA Symposium , 2000. 156–160.

Received 16.03.2020.

Accepted 18.05.2020.

**Аналіз існуючих варіантів класифікації хворих
на серцево-судинними захворюваннями за допомогою нейронними мережами**

Проведено аналіз інформаційних технологій хронобіологічного моніторингу кардіосистем на базі методів класифікації з використанням нейронних мереж. У результаті найкращим класифікатором є нейромережа каскадної кореляції з 85-88% точністю класифікації, найгіршим класифікатором - ймовірна нейронна мережа, оскільки точність цього алгоритму залежить від розміру набору даних.

**Analysis of existing options for the classification of patients
with cardiovascular disease using neural networks**

The work is devoted to the analysis of information technologies of chronobiological monitoring of cardiac systems, development of decision support system for physician-researcher based on classification methods using neural networks such as PNN (Probabilistic Neural Networks), multilayer perceptron and CasCor (Cascade Correlation).

The training sample was 80% of the total number of patients (278 training pairs), and the test sample was 20% (65 training pairs). The presented data sample is highly representative, considering the number of patients: dataset includes 343 patients with cardiovascular complications (class 1 - 142, class 2 - 207). The age range of patients in the sample is 50-60 years old.

The result is the following: the best classifier is the neural network of cascade correlation with 85-88% classification accuracy. The worst classifier was the probabilistic neural network, since the accuracy of this algorithm depends on the size of the data set.

Молодець Богдан Владимирович – асистент, Дніпровський національний університет імені Олеся Гончара.

Буланая Татьяна Михайловна – доцент, к.т.н., Дніпровський національний університет імені Олеся Гончара.

Молодець Богдан Володимирович – асистент, Дніпровський національний університет імені Олеся Гончара.

Булана Тетяна Михайлівна – доцент, к.т.н., Дніпровський національний університет імені Олеся Гончара.

Molodets Bohdan – Assistant Professor, Oles Honchar Dnipro National University.

Bulanaya Tatyana – Associate Professor, Candidate of Technical Sciences (Ph.D.), Oles Honchar Dnipro National University.