

Є. М. Федорченко, А. О. Олійник, С. К. Корнієнко,
О. О. Степаненко, М. В. Саман.

Національний університет «Запорізька політехніка»
вул. Жуковського, 64, 69063 Запоріжжя, Україна

Розробка та дослідження генетичного методу для медичного діагностування цукрового діабету

У результаті дослідження розроблено інтелектуальну систему підтримки прийняття рішень для діагностування цукрового діабету, яка базується на моделях машинного навчання. Розроблено модифікований генетичний метод оптимізації параметрів нейронної мережі. Модифікація простого генетичного алгоритму, яку реалізовано в рамках виконання проєкту, дозволяє пришвидшити виконання підбору параметрів навчання нейронних мереж і підвищити результуючий показник точності порівняно з базовою версією простого генетичного алгоритму, за рахунок модифікації оператора мутації, а також, зміненого підходу до вибору особин для схрещення. Розроблена модель призначена для застосування у сфері медичного обслуговування та дозволяє з певною точністю визначати наявність ризику захворюваності пацієнта на цукровий діабет за клінічними показниками стану здоров'я. Результатом застосування даної моделі є зменшення ймовірності помилки лікаря, підвищення впевненості лікаря в прийнятому рішенні при постановці діагнозу та більша кількість врятованих життів за рахунок постановки вірного та своєчасного діагнозу.

Ключові слова: генетичний алгоритм, класифікація, нейронна мережа, прогнозування, діабет, PSO.

Вступ

На сьогоднішній день існує величезна кількість захворювань, які представляють серйозну небезпеку для здоров'я населення, до них відноситься і цукровий діабет. Вирішення проблеми діагностики діабету є важливою, оскільки хвороба є хронічною та може мати серйозний вплив на весь організм людини.

Цукровий діабет — це глобальна медико-соціальна проблема сучасності. Рівень розповсюдження цукрового діабету зростає в усьому світі. Відповідно до даних Всесвітньої організації охорони здоров'я, розповсюдженість стрімко зростає, наближаючись до глобальної епідемії [1].

Цукровий діабет має такі наслідки: рання інвалідність, високий рівень летальності, заходи, що направлені на реабілітацію хворих, призводять до великих затрат служб охорони здоров'я. Це спричинено тим, що хвороба і її лікування визначаються не з моменту порушення вуглеводного обміну, що визначається лише за допомогою проведення різних тестів навантаження (глюкозо толерантний тест і інші), а при прояві клінічних ознак захворювання, що в свою чергу, є причиною різних судинних та інших змін у тканинах.

Останнім часом велика увага приділяється програмному забезпеченню медичної діяльності. Це стає одним із рушійних факторів розвитку медичної науки, що обумовлює актуальність інженерії і впровадження аналітичних і прогнозуючих програмних систем у медицині. Нині для прогнозування різних захворювань використовують різні методи та моделі, серед яких методи аналізу, часові ряди, регресійні та авторегресійні моделі, методи, що базуються на машинному навчанні — байєсові мережі та штучні нейронні мережі, техніки, що базуються на вирішенні задач фільтрації.

Тож існує актуальність прогнозування захворюваності на цукровий діабет з використанням даних про стан здоров'я людини для планування профілактичних заходів і відповідних фінансових витрат [1].

Аналіз літературних даних і постановка проблеми

Оскільки цукровий діабет є хронічним захворюванням, що важко піддається лікуванню, то його діагностика є актуальним питанням у медичному діагностуванні. Дослідженням даної задачі займається багато науковців по всьому світу, і вже на даний час створено досить багато прогнозних моделей для діагностики та прогнозування цукрового діабету (ЦД), які є як специфічними (придатні лише для прогнозування/діагностування ЦД), так і загальними (придатні для використання для багатьох захворювань).

У роботі [2] представлено модель, що розроблена на основі нейронних мереж і меметичного алгоритму для прогнозування діабету. Дана модель створена на основі демографічних показників способу життя та клінічних характеристик. Для покращення точності прогнозування автори використали меметичний алгоритм для оновлення вагових значень нейронної мережі. Покращивши вагові значення нейромережових моделей, автори отримали модель меметичного алгоритму, яка досягла точності моделі в 93 %. Ця модель здатна забезпечити більшу точність прогнозування, ніж інші моделі, що розглянуті у статті.

Дане дослідження має деякі обмеження. Меметичні алгоритми вимагають порівняно більшого часу на навчання і більш потужний процесор, ніж моделі логістичної регресії. Також кількість даних дещо обмежує якість дослідження. Модель, що представлена в роботі, можна покращити. Точність меметичного алгоритму може бути покращена новими даними та навчанням. Використання цих моделей іншими дослідниками може покращити надійність моделей і, таким чином, зекономити ресурси. В цьому дослідженні меметичні алгоритми були використані для оновлення ваг і покращення точності прогнозування нейронної мережі цукрового діабету.

У роботі [3] пропонується система повноз'язного класифікатора багатошарового перцепторона (БШП), який мінімізує емпіричний ризик і максимізує помил-

ково класифіковані точки в навчальних даних. Дійсно, згорткові нейронні мережі розроблюють множину детекторів ознак і використовують їх для розробки декількох карт, які називаються згортковими шарами. Протягом навчання, мережа визначає, які ознаки є важливими для можливості сканування вхідних даних і для більш точної їхньої класифікації. Виходячи з цього, класифікатор сам розроблює власні особливості детекторів, тобто встановлення зв'язності вхідних сигналів нейрона. В багатьох випадках функції, що розглядаються нейронною мережею, будуть непомітні для людського ока, а саме тому згорткові нейронні мережі найчастіше використовують у задачах прогнозування.

Після реалізації алгоритму в CNN-SVM для класифікації можна зробити висновки, що запропонована система може бути використана для навчання та розробки моделі. Цей аналіз буде включати розробку модифікованої нейронної мережі з глибоким навчання шляхом інтеграції методу опорних векторів. Для того, щоб перевірити систему, показники продуктивності будуть порівнюватися з уже існуючим методом. Запропонована система з модифікованим алгоритмом суттєво оптимізує загальну продуктивність системи визначення і ефективно знижує обчислювальну складність класифікатора.

Автори роботи [4] пропонують модифікацію глибокої навчаючої загорткової нейронної мережі за рахунок інтеграції машини опорних векторів. Традиційна CNN використовує MLP (multilayer perceptron) алгоритми для класифікації похідних ознак від необроблених вхідних даних. Хоча ці алгоритми показали багатообіцяючі результати, MLP-алгоритми мають різноманітні недоліки при класифікації нелінійних наборів даних. Для подолання цих недоліків запропонована архітектура поєднує концепцію CNN (convolutional neural network) методики виділення ознак з класифікатором SVM (support vector machines). Набори функцій з необроблених сигналів обчислюються шляхом математичної згортки цих сигналів з ваговими значеннями ядра Гауса. Розмірність цих мап функцій зменшується з використанням операції максимізації над витягнутими наборами функцій. Ці оптимізовані мапи функцій додатково класифікуються за допомогою SVM з функціями ядра для нелінійної класифікації даних.

Модифікований алгоритм суттєво оптимізує загальну продуктивність системи визначення і ефективно знижує обчислювальну складність класифікатора.

У роботі [5] описаний неінвазивний метод виявлення діабету. Метод включає в себе реалізацію одновимірної згорткової нейронної мережі для класифікації необроблених даних. Алгоритм CNN згортає сигнал необроблених даних за допомогою фільтра на основі ядра (Gaussian Filter Kernel — ядро Гаусова фільтра) для створення набору мап ознак. Розмірність цих мап зменшена шляхом використання операції мінімізації за зразками даних і тим самим знижуючи вибірку ознак. Далі ці мапи застосовуються до повнозв'язного класифікатора, який слідує за механізмом, що подібний до багатошарової нейронної мережі на основі перцептрона.

За допомогою запропонованого авторами роботи [5] алгоритму було виконано успішну класифікацію. Встановлено, що цей алгоритм значно зменшує середньо-квадратичну помилку і оптимізує загальну продуктивність класифікатора, тобто знижує витрати на обчислення та потребу в оптимальних методах розділення функцій.

У роботі [6] запропоновано інтелектуальну прогностичну модель, що використовує глибоке навчання для прогнозування фактора ризику виникнення хвороби у пацієнта та ступені тяжкості діабетиків з використанням набору умовних даних. Модель пропонує глибоке навчання у формі глибокої нейронної мережі, яка допомагає застосовувати прогнозу аналітику до набору даних про діабет для отримання оптимальних результатів. Існуючі прогностичні моделі використовуються для прогнозування ступені тяжкості та факторів ризику діабетиків на основі даних. У запропонованій моделі алгоритм вибору запускається для процесу вибору, далі модуль глибокого навчання має глибоку нейронну мережу, яка використовує обмежену машину Больцмана (RBM — restricted Boltzmann machine) як базової одиниці для аналізу даних, присвоюючи ваги кожній гілці нейронної мережі. Результати тестування показали, що кількість досліджень, проведених по діабету з використанням моделі глибокого навчання, є невеликою. Це допоможе з більшою точністю прогнозувати діабет.

Модель може бути покращена шляхом збільшення розміру даних. Окрім цього, прогнозування може залежати від інших факторів, які відсутні у вказаному наборі даних. Прийняття до уваги цих факторів буде сприяти подальшому підвищенню точності прогнозування.

Точність прогнозування за допомогою цієї моделі може досягати 86 % залежно від набору даних.

Автори роботи [7] побудували прогностичні моделі, використовуючи методи логічної регресії і градієнтного підсилювання (Gradient Boosting Machine — GBM). Площина під кривою робочої характеристики приймача (the area under the receiver operating characteristic curve — AROC) була використана для оцінки дискримінаційної здатності цих моделей. Автори застосували метод скоригованого порогу і метод вагової категорії для покращення чутливості — кількості правильно класифікованих пацієнтів з цукровим діабетом розробленою моделлю. Дані моделі були порівнянні з іншими методами машинного навчання, такими як дерево рішень і випадковий ліс.

Результати моделей показали себе наступним чином: для градієнтного підсилювання чутливість склала 71,6 %, для моделі логічної регресії — 73,4 %.

У роботі [8] автори пропонують прогнозу модель, що має назву покращеного алгоритму прогнозування діабету (Improved Diabetes Prediction Algorithm — IDPA), який поєднує алгоритм ієрархічної кластеризації і алгоритм наївної байєсовської класифікації, що розроблений для виявлення та прогнозування цукрового діабету 2-го типу. В даному випадку модуль IDPA виконує групування даних на дві групи, тобто діабет і недіабет, шляхом застосування алгоритму ієрархічної кластеризації. Далі виконується фільтрація шляхом порівняння значення групи зі значенням класу з наступним застосуванням наївного байєсовського класифікатора для прогнозування діабету. Результати показуються, що запропонований метод може прогнозувати діабет з більш високим рівнем точності (96 %), ніж традиційні та існуючі методи. Дана модель може бути використана для раннього прогнозування діабету, тим самим зменшуючи серйозні ускладнення після нього.

У роботі [9] пропонується гібридна модель для прогнозування діабету 1-го та 2-го типів. Робота моделі складається з двох етапів: автономного (оффлайн) та оперативного (онлайн) етапів. Оффлайн-етап орієнтований на підготовку даних

(структурування) із зовнішніх джерел знань, які представлені у вигляді неструктурованих клінічних діаграм, і онлайн-керівництва по діабету й отримання знань з цих джерел. Отримання знань складається з ручних і автоматичних процедур. При застосуванні цієї моделі клінічні діаграми пацієнтів перетворюються в структуровану форму у вигляді правил. Правила зберігаються в базах знань на оффлайн-процесі. Онлайн-етап — це фаза реального часу або виконання моделі, яка виконує прогнозування і керування лікарів для їхньої підтримки у прийнятті рішення. Ця фаза активується, коли пацієнт прибуває на діагностику або коли пацієнт реєструється для подальшого спостереження. У випадку, якщо пацієнт реєструється вперше, запускається методологія HRBR (hybrid rule-based reasoning). У випадку, коли пацієнт уже зареєстрований, активується лише частина СТА (correlation-based trend analysis). СТА надає можливість лікарям бачити всі попередні записи пацієнта у зведеній формі і виявляти аномальні залежності. СТА також надає аналіз майбутніх тенденцій для всіх спостережень і підтримує їх для визначення потенційних ризиків у майбутньому. Відповідно, лікар може прийняти профілактичні заходи. Точність прогнозування такої моделі складає 95,91 %.

Метод, запропонований у роботі [10], пропонує об'єднання KNN і генетичного алгоритму для підвищення точності класифікації даних діабету.

Автори [10] використали варіації генетичного алгоритму для відбору ознак і ранжування атрибутів, які більше сприяють класифікації. Кращі ранжовані атрибути, що використовуються для прогнозування захворювання і алгоритм класифікації, побудований на основі оцінених атрибутів. Класифікатор навчений класифікувати дані про діабет, які є як діабетичними, так і недіабетичними. Запропонована система відповідальна за: вибір найкращої підмножини ознак з використанням генетичного алгоритму, побудову класифікатора і точності вимірювання класифікатора та виявлення ускладнень, пов'язаних з діабетом, такими як проблеми з серцем і нирками, можливий інсульт та інші.

У роботі [11] запропоновано оптимізовану модель CNN (10-шарову) на основі DCNN та LeNet-5. Дана модель забезпечує точність прогнозування 89–95 % (залежно від набору даних). Було внесено зміни між третім шаром max-pooling та повнозв'язним шаром. Окрім цих шарів CNN також включає три шари згортки, три шари max-pooling та два повнозв'язних шари. Недостатня кількість шарів призводить до неадекватної здатності до вилучення CNN, у той час як їхня надмірна кількість збільшує витрати часу на розрахунок. У даній моделі шари згортки використовуються для вилучення складних параметрів вхідних мап ознак. Шари max-pooling зберігають основні особливості вхідних сигналів при зменшенні параметрів та обчислень, що допомогло уникнути виникнення надмірних розмірів і покращити здатність до узагальнення моделі CNN. Останні два повнозв'язні шари об'єднують усі верхні мапи ознак в одновимірний масив, який використовується для класифікації виходу мережі.

Автори роботи [12] пропонують використовувати сигмоїдне ядро та CNN. Спочатку створена модель SKCNN розраховує сигмоїдне ядро для прогнозування захворювання та препаратів, структурну схожість препаратів і семантичну подібність хвороб. Сигмоїдне ядро препарату поєднується зі структурною схожістю препаратів, а сигмоїдне ядро хвороби поєднується із семантичною схожістю хвороб. Далі CNN для отримання функцій використовує комбіновану схожість пре-

паратів і захворювань. На останньому кроці класифікатор Random Forest вводиться для висновку, чи є препарати та хвороби взаємопов'язаними. Точність, що забезпечує дана модель приблизно становить 92,69 %.

У роботі [13] представлена нова структура CNN, яка полягає у поєднанні ідей двох структур Койчі Іто та Гіла Леві, де Койчі використовував глибоке багатозадачне навчання, а Леві — поділ зображення на п'ять областей. Модифікація CNN, що описана в роботі, складається з двох частин: перша — це чотири загорткових шари з шарами об'єднання; друга частина формується з трьох повнозв'язних шарів. Далі STL та DMTL використовуються в процесі навчання, цей спосіб було визначено Койчі Іто. Точність класифікації даної модифікації становить 80,11 % при використанні CNN та методу навчання STL і 91,34 % — для CNN і методу навчання DMTL.

Математична залежність виникнення цукрового діабету від показників стану та анамнезу пацієнта можна визначити як функцію, в якій можливість виникнення захворювання буде залежати від показників пацієнта. Узагальнена модель залежності має вигляд

$$K_{morb} = f(x_n), \quad (1)$$

де $f()$ — функція залежності; K_{morb} — показник, що характеризує наявність ризику виникнення цукрового діабету; x — сукупність показників, які характеризують стан здоров'я пацієнта; n — кількість показників.

Виходячи з наведених даних і аналізу статистичних даних, можна дійти висновку, що шукана математична модель у нашій роботі буде не детермінованою, а скоріше стохастичною.

Відповідно до Уніфікованого клінічного протоколу первинної і вторинної (спеціалізованої) медичної допомоги [14] оцінка ризику виникнення у пацієнта цукрового діабету проводиться на основі діагностичних заходів та анамнезу людини. До основних факторів розвитку ЦД відносять наступні показники:

- вік пацієнта;
- концентрація глюкози в плазмі через 2 години після її введення в пероральному глюкозотолерантному тесті;
- концентрація інсульту в сироватці крові;
- значення діастолічного артеріального тесту;
- індекс маси тіла пацієнта;
- генетика.

За додаткові показники враховують також кількість вагітностей і товщину шкірної складки.

Окрім вищевказаних факторів впливу на розвиток ЦД є й інші, однак вони не несуть значного впливу на факт виникнення захворювання, тож враховувати їх не має сенсу.

Відповідно до вище вказаних показників здоров'я людини математичну залежність виникнення цукрового діабету можна визначити наступним чином:

$$K_{morb} = f(x_{age}, x_{GC}, x_{IC}, x_{DBP}, x_{BMI}, x_{gen}, x_{NP}, x_{ST}), \quad (2)$$

де x_{age} — вік пацієнта; x_{GC} — концентрація глюкози після її перорального глюкозотолерантного тесту; x_{IC} — концентрація інсуліну; x_{DBP} — значення діастолічно-

го артеріального тесту; x_{BMI} — значення індексу маси тіла; x_{gen} — генетична схильність; x_{NP} — кількість вагітностей; x_{ST} — товщину шкірної складки в районі трицепса.

Мета та задачі дослідження

Об'єкт дослідження — процес прогнозування та виявлення цукрового діабету за медичними показниками стану здоров'я людини.

Предмет дослідження — методи для прогнозування та класифікації захворюваності.

Мета роботи — розробка модифікованого генетичного методу оптимізації параметрів моделі на основі нейронної мережі для вирішення завдання прогнозування захворюваності на цукровий діабет.

Метод дослідження — традиційні моделі, нейронні мережі, комбіновані методи (нейронні мережі та генетичні алгоритми, нейронні мережі та мультиагентні системи).

Розробка модифікованого генетичного методу

Для побудови моделі прогнозування ризику виникнення цукрового діабету як тип штучної нейронної мережі обрано багатошаровий персептрон. Вибір даної нейронної мережі виконано на основі вирішуваної задачі, а саме класифікації даних для визначення наявності ризику виникнення цукрового діабету.

Розроблена нейронна модель має щільну структуру, в якій кожен нейрон зв'язаний з усіма нейронами наступного шару. Нейронна мережа складається із вхідного шару, двох прихованих шарів та одного вихідного нейрона, який визначає кінцевий клас.

Зовнішній шар має 8 нейронів, відповідно до кількості вхідних параметрів, другий шар мережі має 14 нейронів, третій — 7 і вихідний шар має один нейрон.

За функцію активації використано сигмоїдальну функцію, яка фактично видає значення ймовірності ризику захворювання:

$$f(x) = \frac{1}{1 + e^x}, \quad (3)$$

Для первинної ініціалізації синаптичних ваг використано псевдовипадкові значення, що мають нормальний розподіл (середнє значення дорівнює 0, а стандартне відхилення дорівнює 0,05) [14].

Для вирішення поставленого завдання було прийнято рішення розробити модифікацію ГА (генетичний алгоритм), що дозволить підвищити точність нейронної мережі. ГА полягає в паралельній обробці множини рішень, при цьому він фокусується на найбільш перспективних з них.

Рішення задачі оцінюють за допомогою обчислення значень цільової функції. Вигляд цільової функції визначається характером розв'язуваної задачі, для якої використовується генетичний алгоритм. У даному випадку за цільову функцію взято функцію, яка вираховує кількість правильних відповідей нейронної мережі:

$$f = \sum_{i=1}^N \sum_{j=1}^m D(y_{ij} - y_{ij}^t), \quad (4)$$

де m — кількість виходів нейронної мережі; N — кількість зразків у навчальній вибірці; y_{ij} — дійсне значення на j -му виході мережі для образу i ; y_{ij}^t — значення, яке повинно бути на j -му виході мережі для образу i ; D — значення, яке залежить від $|y_{ij} - y_{ij}^t| < d$, якщо нерівність вірна, то $D = 1$, якщо ні, то $D = 0$; d — величина, яка визначає максимально допустиме відхилення реального значення від цільового [15].

У результаті обчислення цільової функції визначається, чи буде продовжено виконання ГА. У випадку продовження виконання ГА виконуються послідовно операції відбору, схрещування та мутації.

Для налаштування ваг нейронної мережі може бути використаний генетичний алгоритм. Для цього необхідно визначити цільову функцію (зазвичай це функція помилки) та спосіб подання вагових значень у хромосомі. Кожна хромосома $K = (k_1, k_2, \dots, k_i)$ — це вектор генів або, в даному випадку, вагових коефіцієнтів (ваги зчитуються з мережі за встановленим порядком — зліва направо та зверху вниз).

Довжина хромосоми відповідно до поставленої задачі оптимізації нейронної мережі вираховується за формулою

$$p = N_1(L+1) + \sum_{\mu=1}^M N_{\mu}(N_{\mu-1} + 1), \quad (5)$$

де N_{μ} — кількість нейронів на μ -му шарі; L — кількість ознак у навчаючій вибірці; M — кількість шарів нейромережі.

Однією із проблем, які виникають під час використання генетичних методів, є визначення не глобального, а локального екстремуму функції, що оптимізується. Для рішення цієї проблеми існує багато модифікацій генетичного методу. В даному проєкті запропоновано обирати батьківські пари на основі їхньої віддаленої спорідненості.

У природі існує два типи схрещування: споріднений і неспоріднений. Аналіз численних матеріалів показує, що ступінь спорідненості має великий вплив на нащадків. Якщо брати приклад з природи, то близька спорідненість супроводжується низькою життєздатністю потомства, має вплив на його пристосованість. При схрещуванні неспоріднених осіб таких проблем не виникає, оскільки шкідливі рецесивні мутації, що знаходяться в гомозиготному стані, перейдуть у гетерозиготний стан і не будуть впливати на життєздатність гібридного організму [16].

Таким чином, на етапі вибору осіб для формування батьківських пар, є сенс враховувати не лише показник пристосованості осіб, але і їхню спорідненість між собою так, щоб пари формувалися з максимально віддалених осіб. Унаслідок цього алгоритм буде переглядати недосліджені області та попереджати збіжність до вже знайденого рішення.

Ступінь спорідненості d_{ij} можна виразити за допомогою формули

$$d_{ij} = \sum_{k=1}^p |x_{ik} - x_{jk}|, \quad (6)$$

де x_{ik} — k -й ген у хромосомі i -го батька; x_{jk} — k -й ген у хромосомі j -го батька; p — довжина хромосоми.

У представлений роботі пропонується два оператори мутації, що відрізняються від класичного тим, що мутація здійснюється не лише для певного відсотка особин, а й певним чином змінюються значення генів у хромосомах. У першому запропонованому методі мутації підлягають 20 % популяції. Далі випадковим чином з хромосоми обираються два гени та вираховується середнє арифметичне цих значень за формулою:

$$H_A = \frac{h_{ij} + h_{ik}}{2}, \quad (7)$$

де h_{ij} — j -й ген i -ї хромосоми; h_{ik} — k -й ген i -ї хромосоми.

Щоб у хромосому не поверталось два гени з однаковими значеннями, пропонується змінювати їхні значення шляхом додавання та віднімання певного числа із заданого діапазону значень:

$$h_{ij}^* = H_A + rand(0,1), \quad (8)$$

$$h_{ik}^* = H_A - rand(0,1), \quad (9)$$

де h_{ij}^* , h_{ik}^* — значення j -го та k -го гена i -ї хромосоми в результаті мутації; $rand(0,1)$ — випадково згенероване число в інтервалі $[0; 1]$.

В якості другої модифікації оператора мутації пропонується обчислення середнього арифметичного значень двох випадково обраних генів (10), але на відміну від першої модифікації, у цій як значення одного гена йтиме значення середнього арифметичного, а як значення другого гена ставимо нуль:

$$h_{ij}^* = H_A + rand(0,1), \quad (10)$$

$$h_{ik}^* = 0, \quad (11)$$

де h_{ij}^* , h_{ik}^* — значення j -го та k -го гена i -ї хромосоми в результаті мутації; $rand(0,1)$ — випадкове значення від нуля до одиниці.

Таким чином, з використанням запропонованих операторів, сенс оператора мутації зберігається, тобто виконується підвищення різноманітності пошуку та введення нових хромосом у популяцію для більш повного дослідження простору пошуку.

Для побудови моделі прогнозування ризику виникнення цукрового діабету за тип штучної нейронної мережі обрано багаточаровий перцептрон. Вибір даної нейронної мережі виконано на основі вирішуваної задачі, а саме класифікації даних для визначення наявності ризику виникнення цукрового діабету.

Після виконання еволюційних операторів оцінюється розмір популяції, після чого її розмір повертається до вихідного. В нову популяцію потрапляють найбільш пристосовані особи.

Окрім цього, також реалізована оптимізація нейронної мережі методом рою часток. Метод рою часток [18] (Particle Swarm Optimization, PSO) являє собою методику стохастичної оптимізації, яка бере ідею у соціальній поведінці зграї птахів або риб.

PSO має багато спільного з генетичними алгоритмами [18]. Система ініціалізується за допомогою популяції випадкових рішень і виконує пошук оптимумів шляхом оновлення поколінь цих рішень. Проте, на відміну від генетичних алгоритмів, PSO не має таких операторів еволюції як кросовер і мутація. У PSO потенційні рішення, що називаються частками, пролітають у просторі рішень, орієнтуючись на частинки, що є оптимальними в кожний момент часу.

У кожній точці, де побувала частка, розраховується значення цільової функції. При цьому кожна частка запам'ятовує, яке (і де) краще значення цільової функції вона особисто знайшла, а також кожна частка знає де розташована точка, яка є кращою серед усіх точок, які розвідали частки. На кожній ітерації частки коректують свою швидкість (модуль і напрямок), щоб з одного боку бути ближче до кращої точки, яку частка знайшла сама (автори алгоритму назвали цей аспект поведінки «ностальгією»), і, в той же час, наблизитися до точки, яка в даний момент є глобально кращою. Через деяку кількість ітерацій частки повинні зібратися поблизу найбільш вдалої точки, хоча можливо, що частина частинок залишиться десь у відносно непоганому локальному екстремуму, але головне, щоб хоча б одна частка виявилася поблизу глобального екстремуму.

У початковому вигляді алгоритму корекція швидкості виглядала наступним чином:

$$v_{i,t+1} = v_{i,t} + \varphi_p r_p (p_i - x_{i,t}) + \varphi_g r_g (g_i - x_{i,t}), \quad (12)$$

де $v_{i,t}$ — i -та компонента швидкості при t -й ітерації алгоритму; $x_{i,t}$ — i -та координата частки при t -й ітерації алгоритму; p_i — i -та координата кращого рішення, знайденого часткою; g_i — i -та координата найкращого рішення, знайденого всіма частками; r_p та r_g — випадкові числа в інтервалі (0, 1); φ_p та φ_g — вагові коефіцієнти, що необхідно підібрати під конкретну задачу.

Після цього коректуємо координату кожної частки:

$$x_{i,t+1} = x_{i,t} + v_{i,t+1}. \quad (13)$$

Далі обчислюємо значення цільової функції у кожній новій точці, кожна частка перевіряє чи не стала нова координата найкращою серед всіх точок, які вона відвідала. Потім серед усіх нових точок перевіряємо, чи не знайдено глобально найкращу точку і, якщо знайдено, запам'ятовуємо її координати та значення цільової функції у цій точці.

Порівняно з генетичними алгоритмами, перевагами PSO є легкість реалізації і невелика кількість параметрів, які треба встановлювати. PSO вдало застосовуються в багатьох областях, таких як оптимізація функцій, тренування штучних нейронних мереж, управління нечіткою системою та інші, де можна застосувати генетичні алгоритми [15].

У загальному вигляді оптимізація нейронної мережі виглядає, як зображено на рис. 1.

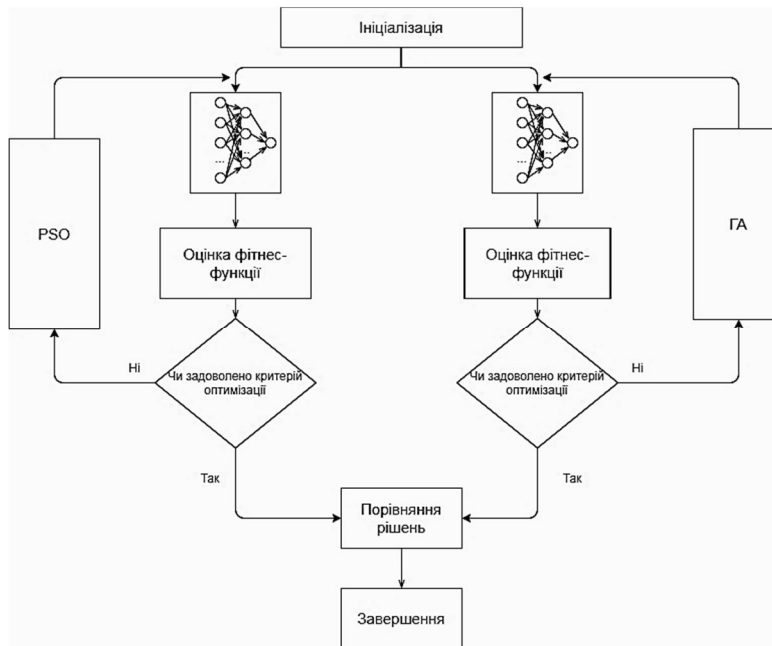


Рис. 1. Оптимізація нейронної мережі за допомогою ГА та PSO

Експерименти навчання методів

Тестування, як і навчання мережі, проводилося на наборі даних, що містять інформацію щодо захворювання на діабет. Джерелом інформації про захворюваність є Національний інститут діабету, захворювань органів травлення та нирок (США) [17]. Набір даних формувався від жінок 21 року і більше народності Піма. Набір складається з 9 атрибутів і 768 екземплярів, а саме:

- 1) кількість вагітностей;
- 2) концентрація глюкози в плазмі через 2 години після введення в пероральному глюкозотолерантному тесті;
- 3) діастолічний артеріальний тиск (мм рт.с.);
- 4) товщина шкірної складки в районі трицепса (мм);
- 5) концентрація інсуліну в сироватці крові (мкЕд/мл);
- 6) індекс маси тіла (вага в кг/(зріст в м)²);
- 7) функція, що описує генетичну схильність до діабету (diabetes pedigree);
- 8) вік (роки);
- 9) клас, що показує чи був у пацієнта ЦД.

Створена мережа проходила навчання впродовж 200 епох. При тестуванні було використано спосіб поділу вибірки даних на навчальну та тестову у відсотковому відношенні 80/20 %.

Правильність класифікації визначається як співвідношення між вірними відповідями та загальною кількістю відповідей:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (14)$$

де TP (True Positive) — випадок, коли реальне значення дорівнювало 1 ($x = 1$) і відповідь алгоритму також 1 ($y = 1$); TN (True Negative) — у здорового пацієнта алгоритм діагностував відсутність хвороби; FP (False Positive) — випадок, коли реальне значення дорівнювало нулю ($x = 0$), а прогнозоване — одиниці ($y = 1$); FN (False Negative) — випадок, коли алгоритм визнав хвору людину здоровою. В даному випадку, оскільки розглядається задача медичної діагностики, $y = 1$ означає наявність хвороби, а $y = 0$ — її відсутність.

Звідси виходить, що в ідеальному випадку модель повинна видавати нуль помилок класів False Positive та False Negative, але насправді така ситуація практично нереальна, тому варто мінімізувати кількість можливих помилок.

Як функція втрат використовується бінарна кросентропія. Бінарна кросентропія обчислює втрати шляхом обчислення наступного середнього:

$$loss = -\frac{1}{output\ size} \sum_{i=1}^{output\ size} y_i \cdot \log \hat{y}_i + (1 - y_i) \cdot \log(1 - \hat{y}_i), \quad (15)$$

де \hat{y}_i — i -та скалярна величина на вході моделі; y_i — відповідне цільове значення; $output\ size$ — кількість скалярних значень даних на виході моделі.

Було використано для оцінки ROC-криву — це графік, що ілюструє ефективність бінарної системи класифікації залежно від заданого класу. ROC-крива надає детальну інформацію про поведінку класифікатора. Крива є результатом зображення на графіку взаємозв'язку True Positive rate (TPR) False Positive rate (FPR).

Ще одним показником, який був використаний для оцінки та порівняння моделей — це час роботи моделі. Час, за який проводиться класифікація може бути свідком того, наскільки швидко можна обробити дані.

Таким чином, тестування розробленого програмного забезпечення має наступні результати.

Звичайний повнозв'язний перцептрон показав такі результати (рис. 2–4). Точність діагностування склала 78,78 %. Графічне відображення процесу навчання моделі та її точність відображені на рис. 2 та в табл. 1.

З рис. 2 можна побачити, що точність діагностування даної моделі склала 78,78 % та спостерігається практично сталі значення протягом навчання повнозв'язного перцептрона.

Процес навчання повнозв'язного перцептрона відповідно до кількості епох відображено в табл. 1, де наведено точність повнозв'язного перцептрона через кожні 25 епох навчання на навчальній і тестовій вибірках.

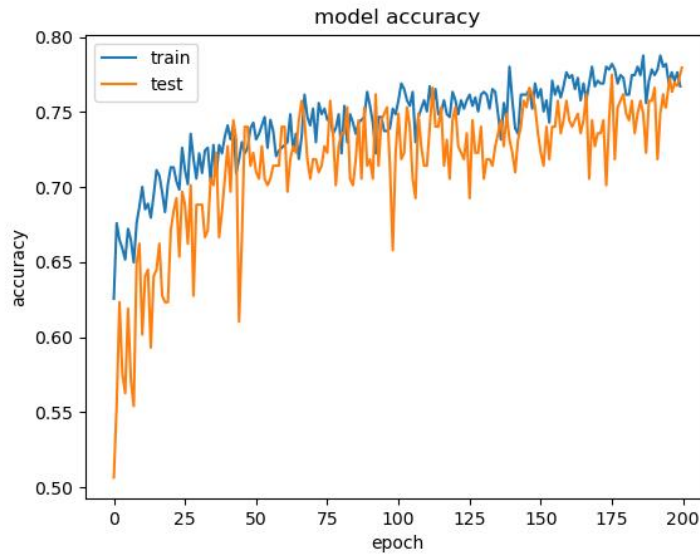


Рис. 2. Точність моделі на основі повнозв'язного перцептрона

Таблиця 1. Точність повнозв'язного перцептрона відповідно до епох навчання

Епоха	0	25	50	75	100	125	150	175	200
Точність на навчальній вибірці, %	62,71	70,77	73,28	74,22	76,15	76,13	76,24	77,76	76,73
Точність на тестовій вибірці, %	51,02	68,85	71,03	72,37	74,94	79,31	72,34	77,29	78,78

Зміну значення помилки відображено на графіку (рис. 3), відповідно до якого спостерігається різке зниження значення помилки, яке проявляє практично стає значення протягом навчання повнозв'язного перцептрона.

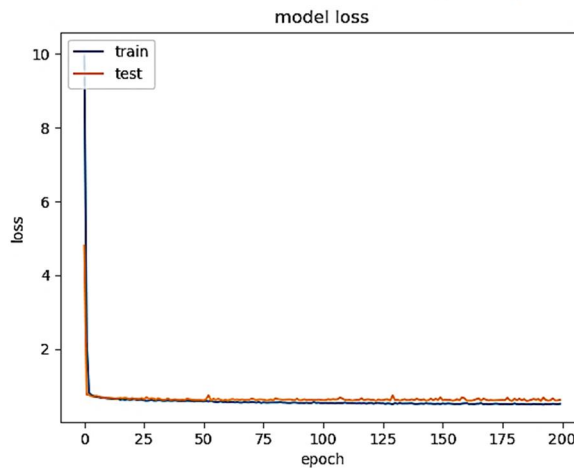


Рис. 3. Оцінка помилки моделі на основі повнозв'язного перцептрона

Процес зниження значення функції втрат повнозв'язного перцептрона в процесі навчання відповідно до кількості епох відображено в табл. 2, де наведено

значення функції втрат повнозв'язного перцептрона через кожні 25 епох навчання на навчальній і тестовій вибірках.

Таблиця 2. Значення функції помилки повнозв'язного перцептрона відповідно до епох навчання

Епоха	0	25	50	75	100	125	150	175	200
Значення функції помилки на навчальній вибірці	10,23	0,641	0,639	0,652	0,649	0,647	0,647	0,647	0,647
Значення функції помилки на тестовій вибірці	10,07	0,61	0,62	0,57	0,59	0,58	0,60	0,59	0,58

Значення ROC-кривої — 0,811, графік якої зображено на рис. 4 та в табл. 3.

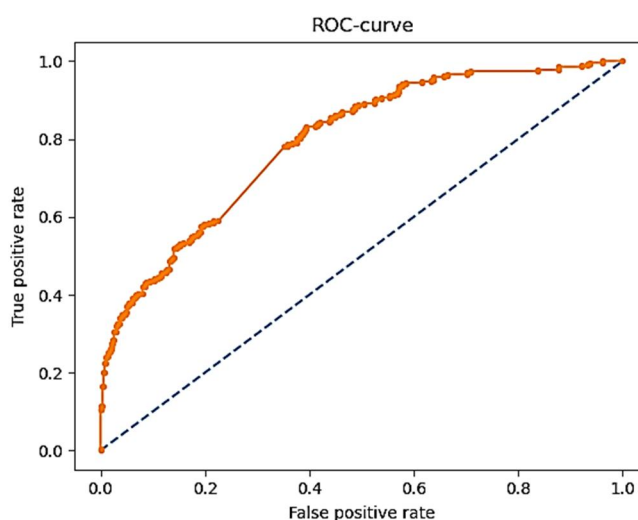


Рис. 4. ROC-крива моделі на основі повнозв'язного перцептрона

Відповідно до збільшення значення False Positive rate відображено зміну значення True Positive rate у табл. 3.

Таблиця 3. Значення ROC-кривої

False Positive rate	0,0	0,2	0,4	0,6	0,8	1,0
True Positive rate	0,0	0,577	0,982	0,944	0,976	1,0

Результати повнозв'язного перцептрона, навчання якого здійснювалося за допомогою ГА зображені на рис. 5–7.

З рис. 5 можна побачити, що точність діагностування даної моделі склала 89,3 %.

Процес навчання повнозв'язного перцептрона за допомогою генетичного алгоритму відповідно до кількості епох відображено в табл. 4, де наведено точність повнозв'язного перцептрона через кожні 25 епох навчання на навчальній і тестовій вибірках.

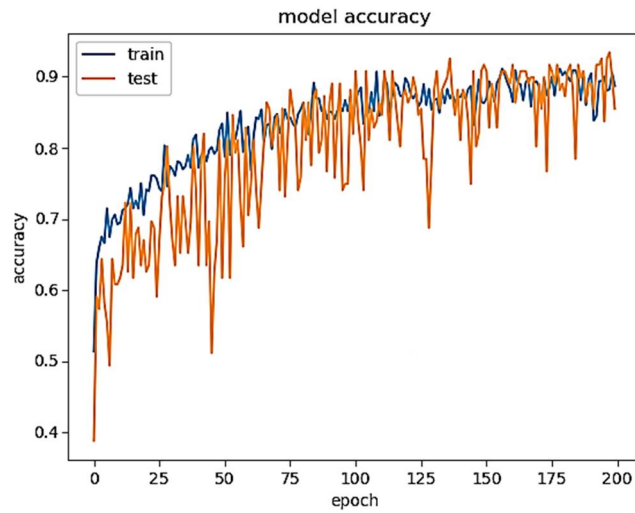


Рис. 5. Точність моделі на основі повнозв'язного перцептрона, що навчалася за допомогою ГА

Таблиця 4. Точність повнозв'язного перцептрона відповідно до епох навчання

Епоха	0	25	50	75	100	125	150	175	200
Точність на навчальній вибірці, %	51,93	74,24	80,47	84,17	87,21	87,34	87,56	89,06	88,87
Точність на тестовій вибірці, %	38,71	67,86	77,12	8,78	90,74	84,34	91,02	89,98	89,3

Зміну значення помилки відображено на графіку (рис. 6), відповідно до якого спостерігається зниження, проте її значення коливається протягом навчання повнозв'язного перцептрона в межах від 0,6 до 1,0.

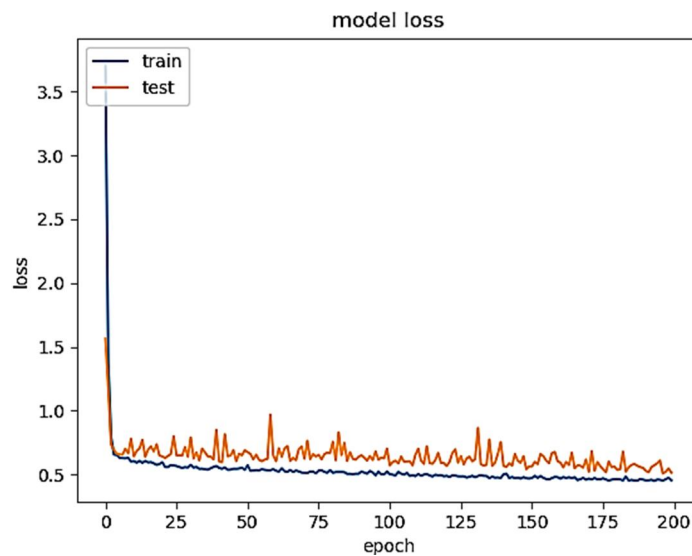


Рис. 6 Оцінка помилки моделі на основі повнозв'язного перцептрона, що навчалася за допомогою ГА

Процес зниження значення функції втрат повнозв'язного перцептрона у процесі навчання генетичним алгоритмом відповідно до кількості епох відображено в табл. 5, де наведено значення функції втрат повнозв'язного перцептрона через кожні 25 епох навчання на навчальній і тестовій вибірках.

Таблиця 5. Значення функції помилки повнозв'язного перцептрона відповідно до епох навчання

Епоха	0	25	50	75	100	125	150	175	200
Значення функції помилки на навчальній вибірці	3,582	0,567	0,551	0,527	0,503	0,497	0,496	0,491	0,487
Значення функції помилки на тестовій вибірці	1,575	0,637	0,671	0,646	0,528	0,530	0,531	0,513	0,511

Графік ROC-кривої зображено на рис. 7 та в табл. 6, значення якої дорівнює 0,884.

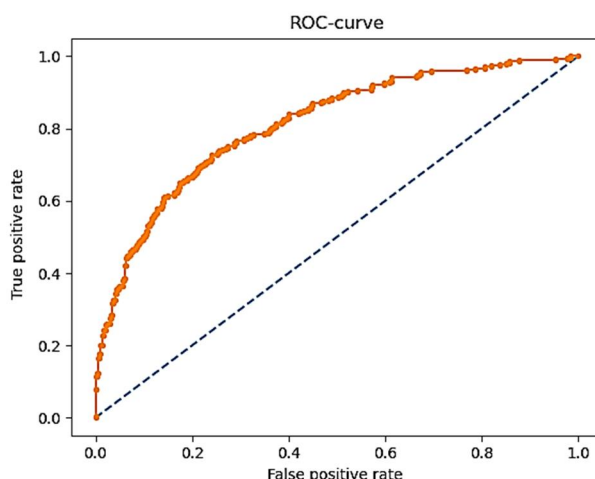


Рис. 7. ROC-крива моделі на основі повнозв'язного перцептрона, що навчалася за допомогою ГА

Відповідно до збільшення значення False Positive rate відображено зміну значення True Positive rate у табл. 6.

Таблиця 6. Значення ROC-кривої

False Positive rate	0,0	0,2	0,4	0,6	0,8	1,0
True Positive rate	0,0	0,681	0,843	0,928	0,973	1,0

Результати повнозв'язного перцептрона, навчання якого здійснювалося за допомогою PSO зображені на рис. 8–10.

З рис. 8 можна побачити, що точність діагностування повнозв'язного перцептрона, навчання якого здійснювалося за допомогою PSO, склала 90,82 %.

Процес навчання повнозв'язного перцептрона за допомогою методу рою часток відповідно до кількості епох відображено в табл. 7, де наведено точність повнозв'язного перцептрона через кожні 25 епох навчання на навчальній і тестовій вибірках.

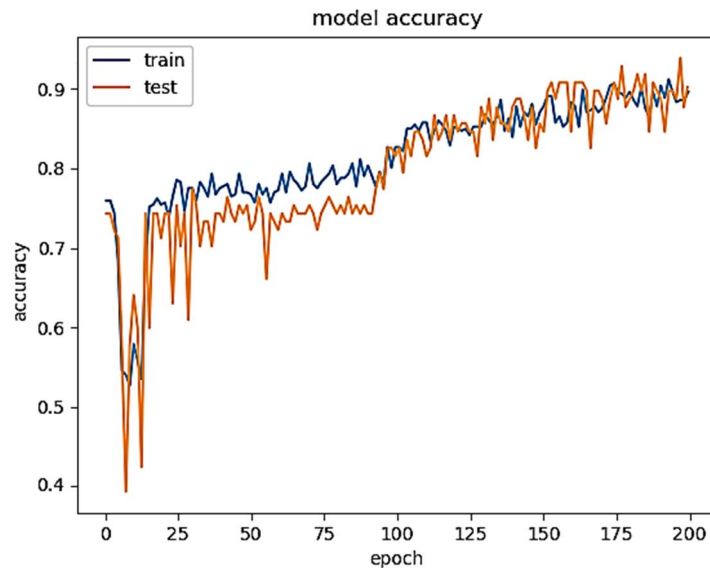


Рис. 8. Точність моделі на основі повнозв'язного перцептрона, що навчалася за допомогою PSO

Таблиця 7. Точність повнозв'язного перцептрона відповідно до епох навчання

Епоха	0	25	50	75	100	125	150	175	200
Точність на навчальній вибірці, %	76,23	78,35	76,57	78,59	82,78	84,76	87,88	90,43	89,94
Точність на тестовій вибірці, %	74,49	73,05	73,31	75,47	82,38	84,53	89,55	89,27	90,82

Зміну значення помилки відображено на графіку (рис. 9), відповідно до якого спостерігається зниження, проте її значення коливається протягом навчання мережі в межах від 0,5 до 1,0.

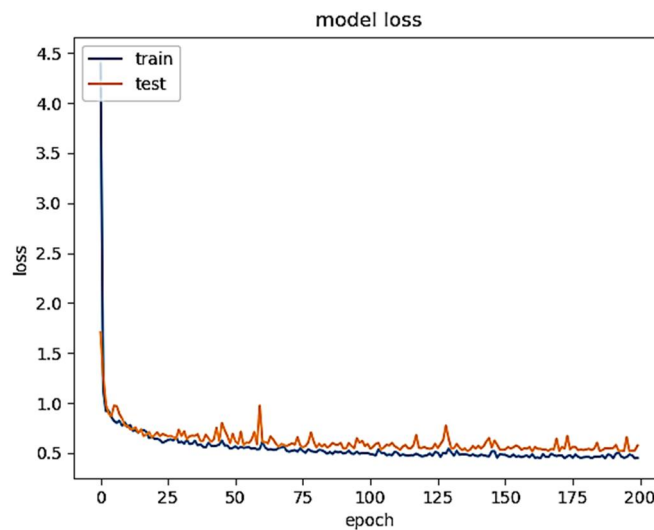


Рис. 9. Оцінка помилки моделі на основі повнозв'язного перцептрона, що навчалася за допомогою PSO

Процес зниження значення функції втрат повнозв'язного перцептрона, що навчався за допомогою методу рою часток у процесі навчання відповідно до кількості епох, відображено в табл. 8, де наведено значення функції втрат повнозв'язного перцептрона через кожні 25 епох навчання на навчальній і тестовій вибірках.

Таблиця 8. Значення функції помилки повнозв'язного перцептрона відповідно до епох навчання

Епоха	0	25	50	75	100	125	150	175	200
Значення функції помилки на навчальній вибірці	4,497	0,637	0,552	0,537	0,498	0,501	0,509	0,510	0,489
Значення функції помилки на тестовій вибірці	1,72	0,69	0,575	0,568	0,597	0,566	0,567	0,566	0,564

Графік ROC-кривої наведено на рис. 10 та в табл. 9. Її значення складо 0,893.

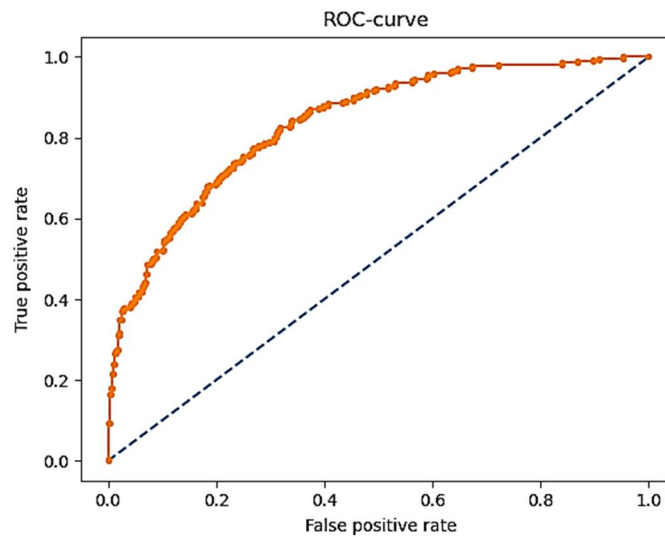


Рис. 10. ROC-крива моделі на основі повнозв'язного перцептрона, що навчалася за допомогою PSO

Відповідно до збільшення значення False Positive rate відображено зміну значення True Positive rate у табл. 9.

Таблиця 9. Значення ROC-кривої

False Positive rate	0.0	0.2	0.4	0.6	0.8	1.0
True Positive rate	0,0	0,67	0,891	0,965	0,983	1,0

Результати моделі на основі повнозв'язного перцептрона, навчання якої здійснювалося за допомогою ГА з модифікованими операторами селекції та з першим із запропонованих операторів мутації, зображені на рис. 11–13, графік точності моделі наведено на рис. 11 та в табл. 10.

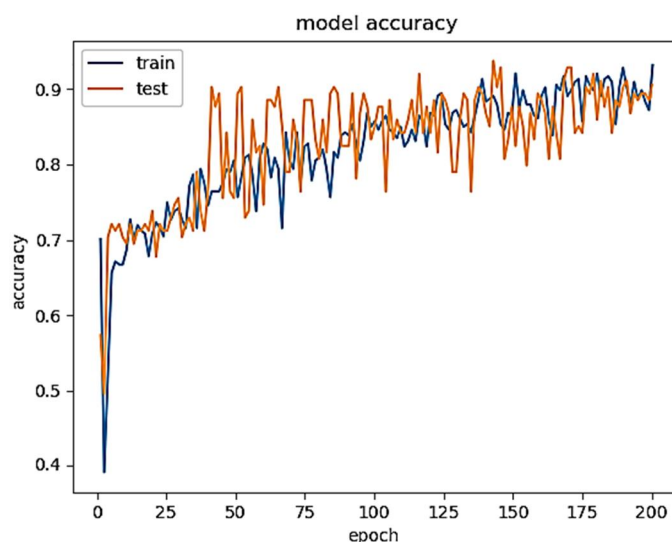


Рис. 11. Точність нейронної мережі, що навчалася за допомогою першої модифікації ГА

З рис. 11 можна побачити, що точність діагностування повнозв'язного перцептрона, навчання якої здійснювалося за допомогою ГА з модифікованими операторами селекції і з першим із запропонованих операторів мутації, склала 91,3 %.

Процес навчання повнозв'язного перцептрона за допомогою генетичного алгоритму із запропонованим оператором відбору та першого підходу до оператору мутації відповідно до кількості епох відображено в табл. 10, де наведено точність повнозв'язного перцептрона через кожні 25 епох навчання на навчальній і тестовій вибірках.

Таблиця 10. Точність повнозв'язного перцептрона відповідно до епох навчання

Епоха	0	25	50	75	100	125	150	175	200
Точність на навчальній вибірці, %	70,58	73,85	77,56	82,53	85,37	85,41	88,90	89,12	83,26
Точність на тестовій вибірці, %	57,78	71,23	84,79	88,62	85,10	88,74	85,98	85,37	91,3

Зміну значення помилки відображено на графіку (рис. 12), відповідно до якого спостерігається зниження значення з незначними коливаннями в межах значень 0,5–0,75.

Процес зниження значення функції втрат повнозв'язного перцептрона, що навчався за допомогою генетичного алгоритму із запропонованим оператором відбору та першого підходу до оператору мутації у процесі навчання відповідно до кількості епох, відображено в табл. 11, де наведено значення функції втрат повнозв'язного перцептрона через кожні 25 епох навчання на навчальній і тестовій вибірках.

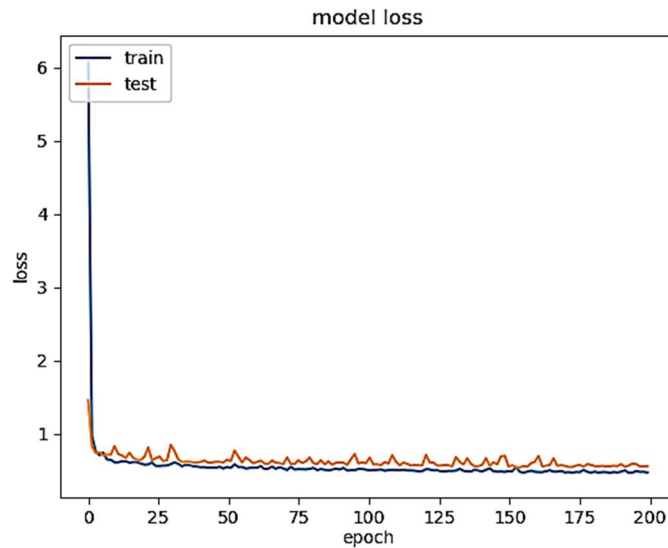


Рис. 12. Оцінка помилки моделі на основі повнозв'язного перцептрона, що навчалася за допомогою першої модифікації ГА

Таблиця 11. Значення функції помилки повнозв'язного перцептрона відповідно до епох навчання

Епоха	0	25	50	75	100	125	150	175	200
Значення функції помилки на навчальній вибірці	6,975	0,563	0,471	0,468	0,403	0,407	0,410	0,406	0,404
Значення функції помилки на тестовій вибірці	1,50	0,679	0,619	0,617	0,615	0,613	0,612	0,611	0,610

Значення ROC-кривої складає 0,9, графік якої зображено на рис. 13 та в табл. 12.

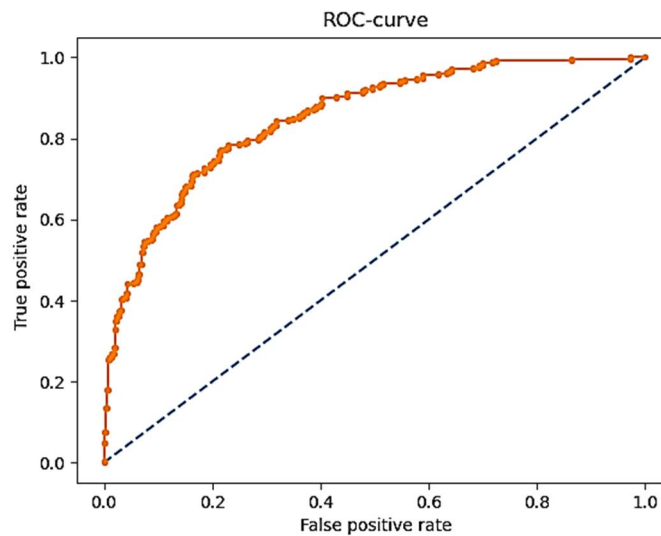


Рис. 13. ROC-крива моделі на основі повнозв'язного перцептрона, що навчалася за допомогою першої модифікації ГА

Відповідно до збільшення значення False Positive rate відображено зміну значення True Positive rate у табл. 12.

Таблиця 12. Значення ROC-кривої

False Positive rate	0,0	0,2	0,4	0,6	0,8	1,0
True Positive rate	0,0	0,689	0,896	0,961	0,985	1,0

Результати повнозв'язного перцептрона, навчання якого здійснювалося за допомогою ГА з модифікованими операторами селекції і з другим із запропонованих операторів мутації, зображені на рис. 14–16.

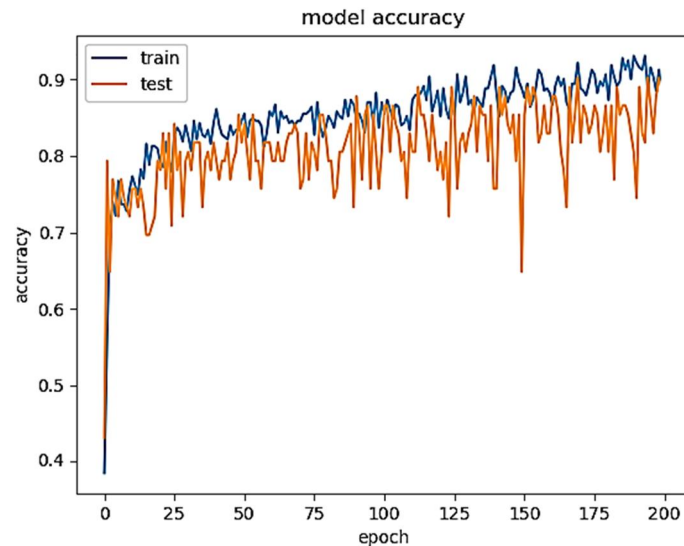


Рис. 14. Точність моделі на основі повнозв'язного перцептрона, що навчалася за допомогою другої модифікації ГА

З рис. 14 можна побачити, що точність діагностування повнозв'язного перцептрона, навчання якого здійснювалося за допомогою ГА з модифікованими операторами селекції та з другим із запропонованих операторів мутації склала 89,97 %.

Процес навчання повнозв'язного перцептрона за допомогою генетичного алгоритму із запропонованим оператором відбору та другого підходу до оператора мутації відповідно до кількості епох відображено в табл. 13, де наведено точність повнозв'язного перцептрона через кожні 25 епох навчання на навчальній та тестовій вибірках.

Таблиця 13. Точність повнозв'язного перцептрона відповідно до епох навчання

Епоха	0	25	50	75	100	125	150	175	200
Точність на навчальній вибірці, %	36,74	82,94	84,52	82,98	86,69	87,43	87,85	90,38	89,14
Точність на тестовій вибірці, %	43,31	83,74	82,76	80,73	86,65	82,83	85,56	86,67	89,97

Зміну значення помилки відображено на графіку (рис. 15) та в табл. 14, відповідно до якого спостерігається поступове зниження значення помилки, що зберігає стабільне значення, починаючи з 75 епохи.

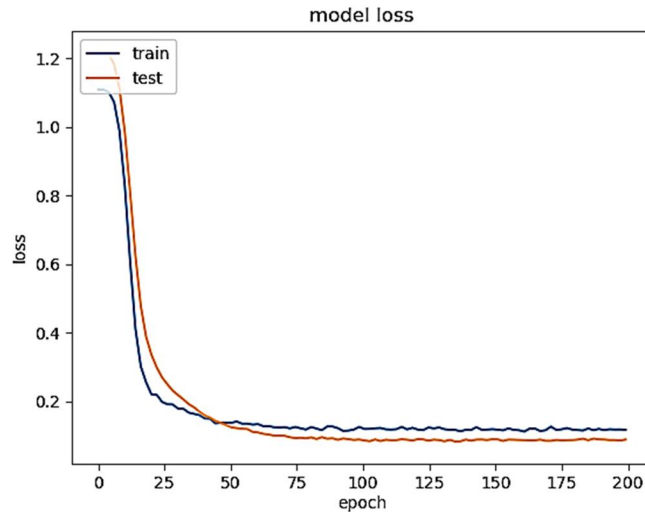


Рис. 15. Оцінка помилки повнозв'язного перцептрона, що навчалася за допомогою другої модифікації ГА

Таблиця 14. Значення функції помилки повнозв'язного перцептрона відповідно до епох навчання

Епоха	0	25	50	75	100	125	150	175	200
Значення функції помилки на навчальній вибірці	1,157	0,198	0,146	0,125	0,123	0,124	0,123	0,124	0,123
Значення функції помилки на тестовій вибірці	1,33	0,264	0,132	0,088	0,081	0,080	0,081	0,080	0,080

Процес зниження значення функції втрат моделі на основі повнозв'язного перцептрона, що навчалася за допомогою генетичного алгоритму із запропонованим оператором відбору та першого підходу до оператору мутації у процесі навчання відповідно до кількості епох, відображено в табл. 14, де наведено значення функції втрат повнозв'язного перцептрона через кожні 25 епох навчання на навчальній і тестовій вибірках.

Значення ROC-кривої дорівнює 0,899. Графік даної кривої зображено на рис. 16 та в табл. 15.

Відповідно до збільшення значення False Positive rate відображено зміну значення True Positive rate у табл. 15.

Для наочного відображення результатів, що отримані в процесі дослідження моделей, дані занесені до табл. 16, а точність класифікації відображена на гістограмі (рис. 17).

Результати порівняння значення точності всіх експериментів відображено на гістограмі (рис. 17). Видно, що навчання нейронної мережі за допомогою еволюційних методів суттєво збільшують їхню точність як класифікатора.

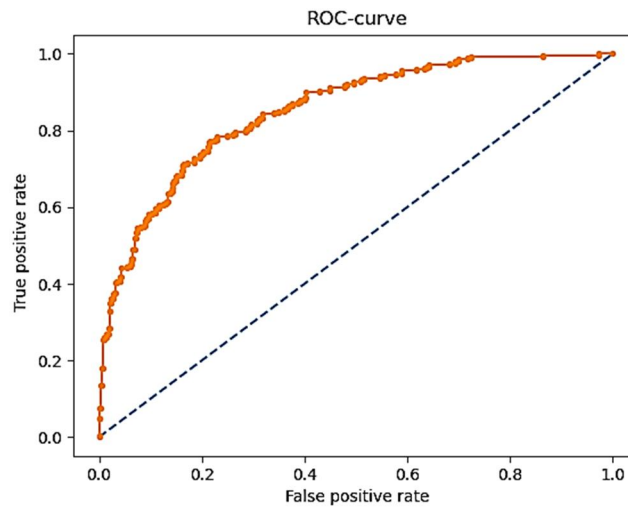


Рис. 16. ROC-крива моделі на основі повнозв'язного перцептрона, що навчалася за допомогою другої модифікації ГА

Таблиця 15. Значення ROC-кривої

False Positive rate	0,0	0,2	0,4	0,6	0,8	1,0
True Positive rate	0,0	0,743	0,872	0,957	0,983	1,0

Таблиця 16. Порівняльна таблиця результатів оцінювання моделей

№ п/п	Назва	Точність, %	Значення ROC-кривої	Час, с
1	Нейронна мережа	78,78	0,811	117
2	Нейронна мережі + ГА	89,3	0,884	99
3	Нейронна мережі + перша модифікація ГА	91,3	0,9	101
4	Нейронна мережі + друга модифікація ГА	89,97	0,899	102
5	Нейронна мережі + PSO	90,82	0,893	97

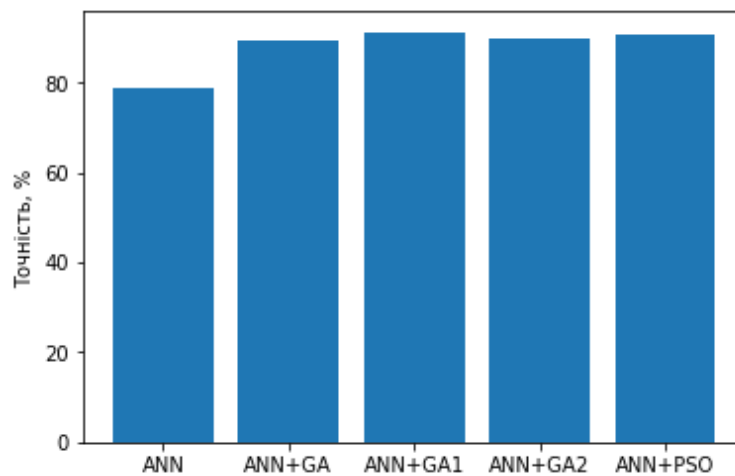


Рис. 17. Гістограма порівняння точності тестованих моделей

Обговорення результатів

З результатів проведених експериментів можна зробити висновок про те, що навчання мережі за допомогою еволюційних алгоритмів показує значне покращення порівняно зі стандартними методами. Використання розглянутих методів показує приблизно однаковий результат за точністю діагностування, а час, що було затрачено на роботу, є доволі незначним.

З рис. 17 видно, що найкращу точність показав повнозв'язний перцептрон, що навчався за допомогою генетичного алгоритму із запропонованим оператором відбору та першого підходу до оператора мутації, точність якого складала 91,3 %. Повнозв'язний перцептрон, що навчався за допомогою методу рою часток показав точність у 90,82 %. Трохи гірші результати має повнозв'язний перцептрон, що навчався за допомогою генетичного алгоритму із запропонованим оператором відбору та другого підходу до оператора мутації, точність якого складала 89,97 %. Повнозв'язний перцептрон, що навчався за допомогою генетичного алгоритму показав точність у 89,3 %. Найгірший результат за точністю має повнозв'язний перцептрон, що навчався за допомогою градієнтного методу, його точність складала 78,78 %.

Висновки

Під час виконання роботи було проаналізовано класичні методи, що можуть бути застосовані для задачі класифікації, а саме Байєсова мережа, NNge, Random Forest, HMM, RBFNetwork, AdaBoost та ENORA. Визначено, що вони не видають задовільної точності при класифікації даних. Найкращим чином себе показали Байєсова мережа та алгоритм Random Forest. Також для порівняння було проаналізовано багатосаровий перцептрон, згорткову нейронну мережу, мережу LSTM та повнозв'язну НМ. Найкращу точність продемонструвала мережа LSTM, але для задачі класифікації вона має доволі складну структуру.

У результаті виконання роботи запропоновано спосіб навчання нейронної мережі за допомогою генетичного алгоритму, для якого запропоновано спосіб вибору осіб до батьківських пар на основі їхньої неспорідненості та два способи мутації генів хромосом. Запропоновано два підходи до оператора мутації. При мутації значення обраних генів обчислюються на основі їхніх середніх значень, після чого додається, віднімається або обнуляється значення.

Під час створення та тренування моделі на основі мережі довгої короткочасної пам'яті було досліджено можливість використання методу рою часток і генетичного алгоритму для налаштування ваг мережі. За результатами тестування та проведення експериментального дослідження визначено, що еволюційні методи значно підвищують точність класифікації. Точність класифікації повнозв'язного перцептрона складала 78,78 %. Точність класифікації повнозв'язного перцептрона, що навчався за допомогою генетичного алгоритму складала 89,3 %, за допомогою PSO — 90,82 %. Повнозв'язний перцептрон, що навчався за допомогою генетичного алгоритму із запропонованими операторами вибору та першого підходу до мутації показав точність у 91,3 %, а з другим підходом до виконання мутації точність класифікатора складала 89,97 %. За результатами досліджень запропонованого генетичного алгоритму можна сказати, що спостерігається поліпшення за точністю класифікації, а саме у першому випадку порівняно із класичним ГА на 2 %, в

другому випадку на 0,67 %. Отримані результати дозволяють запропонувати ефективні методи для підвищення якості прогнозування наявності цукрового діабету, що поліпшить процес діагностики та лікування захворювання.

1. Parte R.S., Patil A., Patil A., Kad A., Kharat S. Non-invasive method for diabetes detection using CNN and SVM classifier. *International journal of research in engineering, science and management*. 2019. Vol. 2. P. 659–661. ISSN: 2581–7175.
2. Jahani M., Mahdavi M. Comparison of Predictive Models for the Early Diagnosis of Diabetes. *Healthcare Informatics Research*. 2016. Vol. 22. P. 95–100. doi: 10.4258/hir.2016.22.2.95.
3. Farre P., Heurteau A., Cuvier O., Emberly E. Dense neural networks for predicting chromatin conformation. *BMC Bioinformatics*. 2018. Vol. 19. doi: 10.1186/s12859-018-2286-z.
4. Lekha S., Suchetha M. A Novel 1-D convolution neural network with SVM architecture for real-time detection applications. *IEEE Sensors Journal*. 2018. Vol. 18. P. 724–731. doi: 10.1109/jsen.2017.2780178.
5. Lekha S., Suchetha M. Real-Time Non-Invasive Detection and Classification of Diabetes Using Modified Convolution Neural Network. *IEEE Journal of Biomedical and Health Informatics*. 2018. Vol. 22. P. 1630–1636. doi: 10.1109/jbhi.2017.2757510.
6. Balaji H., Iyengar N., Caytiles R. Optimal Predictive analytics of Pima Diabetics using Deep Learning. *International Journal of Database Theory and Application*. 2017. Vol. 10. P. 47–62. doi: 10.14257/ijda.2017.10.9.05.
7. Lai H., Huang H., Keshavjee K., Guergachi A., Gao X. Predictive models for diabetes mellitus using machine learning techniques. *BMC Endocrine Disorders*. 2019. Vol. 19. doi: 10.1186/s12902-019-0436-6.
8. Nimmagadda S.P., Yeruva S., Siempu R. Improved diabetes prediction model for predicting Type-II diabetes. *International journal of innovative technology and exploring engineering*. 2019. Vol. 8. P. 230–235. doi: 10.35940/ijitee.l3594.1081219.
9. Ali R., Hussain J., Siddiqi M., Hussain M., Lee S. H2RM: A Hybrid Rough Set Reasoning Model for Prediction and Management of Diabetes Mellitus. *Sensors*. 2015. Vol. 15. P. 15921–15951. doi: 10.3390/s150715921.
10. Pavate A., Ansari N. Risk Prediction of Disease Complications in Type 2 Diabetes Patients Using Soft Computing Techniques. 2015 Fifth International Conference on Advances in Computing and Communications (ICACC). 2015. doi: 10.1109/icacc.2015.61.
11. Li G., Watanabe K., Anzai H., Song X., Qiao A., Ohta M. Pulse-Wave-Pattern Classification with a Convolutional Neural Network. *Scientific Reports*. 2019. Vol. 9. Doi: 10.1038/s41598-019-51334-2.
12. Jiang H., You Z., Huang Y. Predicting drug-disease associations via sigmoid kernel-based convolutional neural networks. *Journal of Translational Medicine*. 2019. Vol. 17. doi: 10.1186/s12967-019-2127-5.
13. Al-Azzawi D. Human Age and Gender Prediction Using Deep Multi-Task Convolutional Neural Network. *Journal of Southwest Jiaotong University*. 2019. Vol. 54. doi: 10.35741/issn.0258-2724.54.4.11.
14. Hemanth D., Anitha J., Naaji A., Geman O., Popescu D., Hoang Son L. A Modified Deep Convolutional Neural Network for Abnormal Brain Image Classification. *IEEE Access*. 2019. Vol. 7. P. 4275–4283. doi: 10.1109/access.2018.2885639.
15. Panda M. Combining multi-objective evolutionary algorithm with averaged one-dependence estimators for big data analytics. *International Journal of Computational Intelligence Studies*. 2018. Vol. 7. P. 1. doi: 10.1504/ijcistudies.2018.090160.
16. Sisodia D., Sisodia D. Prediction of Diabetes using Classification Algorithms. *Procedia Computer Science*. 2018. Vol. 132. P. 1578–1585. doi: 10.1016/j.procs.2018.05.122.
17. Zou Q., Qu K., Luo Y., Yin D., Ju Y., Tang H. Predicting Diabetes Mellitus With Machine Learning Techniques. *Frontiers in Genetics*. 2018. Vol. 9. doi: 10.3389/fgene.2018.00515.
18. Kalyankar G., Poojara S., Dharwadkar N. Predictive analysis of diabetic patient data using machine learning and Hadoop. 2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC). 2017. doi: 10.1109/i-smac.2017.8058253.

Надійшла до редакції 05.06.2021