



Р. О. Ткаченко, І. В. Ізонін, В. М. Данилик, В. Ю. Михалевич

Національний університет "Львівська політехніка", м. Львів, Україна

СТЕКІНГ НЕЙРОПОДІБНОЇ СТРУКТУРИ МПГП З RBF ШАРОМ НА ПІДСТАВІ ГЕНЕРУВАННЯ ВИПАДКОВОГО КОРТЕЖУ ЇЇ ГІПЕРПАРАМЕТРІВ ДЛЯ ЗАВДАНЬ ПРОГНОЗУВАННЯ

Підвищення точності прогнозування засобами штучного інтелекту є важливим завданням в різних галузях промисловості, економіки, медицини. Ансамблеве навчання є одним із можливих варіантів досягнення цього. Зокрема, побудова стекінгових моделей на підставі різних методів машинного навчання, чи з використанням різних частин наявного набору даних демонструє високу точність прогнозу. Проте потребу правильного підбору членів ансамблю, їх оптимальних параметрів, тощо, зумовлює потребу великих часових витрат на підготовку та навчання таких моделей. В роботі пропонується дещо інший підхід до побудови простого, проте ефективного ансамблевого методу. Розроблено нову модель стекінгу нелінійних нейроподібних структур МПГП, яка базується на використанні тільки одного типу ШНМ як елементної бази ансамблю та застосуванні однакової для усіх членів ансамблю навчальної вибірки. Такий підхід забезпечує ряд переваг порівняно з процедурами побудови ансамблів на підставі різних методів машинного навчання, як мінімум у напрямі підбору оптимальних параметрів для кожного з них. За основу ансамлювання в нашому випадку використано кортеж випадкових гіперпараметрів для кожного окремого члена ансамблю. Тобто навчання кожної комбінованої нейроподібної структури МПГП з додатковим RBF шаром, як окремого члена ансамблю відбувається з використанням різних, випадково обраних значеннях центрів RBF та центрів мас. Це забезпечує необхідне різноманіття елементів ансамблю. Експериментальні дослідження щодо ефективності роботи розробленого ансамблю проведено із використанням реального набору даних. Задача полягає у прогнозуванні величини медичних страхових виплат на підставі ряду незалежних атрибутів. Експериментальним шляхом визначено оптимальну кількість членів ансамблю, яка забезпечує найвищу точність розв'язання поставленої задачі. Проведено порівняння результатів роботи розробленого ансамблю з наявними методами цього класу. Встановлено найвищу точність роботи розробленого ансамблю при задовільній тривалості процедури його навчання.

Ключові слова: стекінг; ансамблеве навчання; нейроподібні структури; радіально-базисні функції; медичні страхові виплати.

Вступ

Розв'язок складних завдань прогнозування супроводжується необхідністю розроблення нових, більш ефективних методів машинного навчання [2], [6]. У випадках, коли обмеженнями на обчислювальні ресурси можна нівелювати, а точність прогнозу являється першим пріоритетом, варто застосовувати ансамблі моделей [16], [17]. Здебільшого такі засоби демонструють істотне збільшення точності прогнозу порівняно з батьківськими алгоритмом машинного навчання чи моделлю нейронної мережі [5], [12], [25].

Ідея побудови ансамблевих методів спирається на так званий принцип "мудрість натовпу". Він полягає у можливості врахування великої кількості різних "думок" з метою підвищення точності формування остаточного результату. При правильній постановці задачі та при правильних внутрішніх зв'язках в середині "натовпу", такий підхід демонструє ефективні результати. Забезпечення подібного різноманіття "думок" в стекінговому варіанті ансамблевої реалізації цього принципу може досягатися за рахунок:

- 1) поєднання різних методів машинного навчання чи штучних нейронних мереж (або і тих і інших) як елементної бази ансамблю;
- 2) використання різних частин однієї вибірки (з повторами чи без повторень).

Очевидно, що використання першого підходу супроводжується великою кількістю необхідних ресурсів для навчання, налагодження та тестування ансамблевої моделі. Окрім цього, поєднання різних інструментів машинного навчання в одну ансамблеву модель вимагає чимало часу для підбору оптимальних параметрів роботи кожного окремого члена ансамблю. Використання різних частин вибірки вимагає наявності великої кількості векторів для навчання. Проте такі дані не завжди доступні. Тому виникає задача побудови простої і ефективної ансамблевої моделі яка нівелюватиме вищевказані обмеження.

Об'єкт дослідження – прогнозування нейроподібних засобами машинного навчання.

Предмет дослідження – ансамблі штучних нейроподібних структур з неітеративним машинним навчанням.

Мета роботи – побудова стекінгового ансамблю на підставі використання неітеративної нелінійної нейронної мережі одного типу, яка забезпечуватиме високі показники точності прогнозу.

Для досягнення зазначеної мети визначено такі основні завдання дослідження:

- розроблення стекінгового ансамблю на підставі неітеративних засобів штучного інтелекту;
- підбір оптимальних параметрів його роботи; та апробацію роботи ансамблевого методу під час розв'язання задачі прогнозування.

Наукова новизна отриманих результатів дослідження – вперше розроблено модель стекингу нелінійних нейроподібних структур МПГП, яка базується на використанні тільки одного типу ШНМ як елементної бази ансамблю та застосуванні однакової для усіх членів ансамблю навчальної вибірки. Різноманіття елементів ансамблю в цьому випадку забезпечується за рахунок генерування кортежу випадкових гіперпараметрів для кожного окремого члена ансамблю.

Практична значущість результатів дослідження – простота побудови моделі за рахунок використання тільки одного типу ШНМ та використання загального для усіх членів ансамблю набору даних; висока швидкість роботи моделі за рахунок неітеративного характеру процедури навчання комбінованої нейроподібної структури МПГП з додатковим RBF шаром; уникнення потреби налаштування великої кількості гіперпараметрів окремих моделей ансамблю за рахунок використання ШНМ тільки одного типу як елементної бази ансамблю; підвищення точності прогнозу за рахунок додаткового RBF-шару нейроподібних структур МПГП; наявність екстраполятивних властивостей, що є перевагою комбінованої нейроподібної структури МПГП з додатковим RBF шаром перед RBF нейронною мережею та її різновидами.

Аналіз останніх досліджень та публікацій. Серед різноманіття парадигм машинного навчання, які активно розвиваються в останні роки, особливе місце посідає ансамблеве навчання. Парадигма ансамблевого навчання передбачає об'єднання кількох моделей для формування більш точного результату розв'язання поставленої задачі.

Серед основних класів методів ансамблювання виділяють бегінг, бустинг і стекинг [13]. Ідея методів першого класу полягає у навчанні слабких однорідних класифікаторів та їх об'єднанні за деяким визначеним правилом усереднення. Методи другого класу передбачають навчання окремих однорідних моделей послідовно, де кожна наступна враховує помилки попередньої, та їх об'єднання знову ж таки за визначеним правилом. Третій клас методів вимагає виконання паралельної процедури навчання різнорідних моделей, вихідні сигнали роботи яких усереднюються, або передаються певному метаалгоритму для формування остаточного результату.

З точки зору врахування "думки" різних моделей, а також можливості їх паралельної роботи, що істотно зменшує час навчання моделі, стекингові методи виглядають найбільш привабливо.

У роботі [9] розроблено стекингову модель лінійного жадібного алгоритму для вирішення завдань класифікації в області медицини. Розроблений метод доволі швидкий, та легко піддається інтерпретації. Він забезпечує значно вищу точність роботи порівняно з рядом окремих лінійних моделей. Проте лінійні алгоритми, що утворюють ансамбль не завжди забезпечать високу точність роботи, що накладає ряд обмежень на його використання.

Роботу [3] присвячено побудові стекингового ансамблю з використанням машини опорних векторів, k -найближчих сусідів, наївного Баєсового класифікатора та багатопарового перцептронну. Як метаалгоритм обрали логістичну регресію. У роботі [11] розроблено подібну, проте багаторівневу стекингову модель на підставі

алгоритмів машинного навчання. Серед переваг розроблених методів варто відзначити високу точність роботи порівняно з наявними методами. Проте багаторівневе представлення, а також потребу налаштування кожного окремого елемента ансамблю вимагає чималих часових ресурсів, а також високої кваліфікації оператора.

У роботі [1] розроблено адаптивну модель стекингу. Різноманіття елементів моделі тут досягається за рахунок створення ансамблю несуміжних базових класифікаторів. Кожен з елементів такого ансамблю навчається на підставі своєї унікальної частини із доступного набору даних. Авторами встановлено значне підвищення точності роботи розробленого ансамблю як на збалансованих так і незбалансованих наборах даних. Проте реалізація такого підходу вимагає великої навчальної вибірки, що не завжди можливо.

Матеріали і методи дослідження. Комбінована нейроподібна структура МПГП з RBF розширенням входів. У роботі [20] розроблено дві модифікації неітеративної лінійної нейроподібної структури Моделі Послідовних Геометричних Перетворень (МПГП). Обидві базуються на використанні додаткового шару на підставі RBF-функцій. Основне його призначення це нелінійне розширення початкових входів задачі з метою отримання більш точних результатів прогнозу. Детальні математичні викладки процедур навчання і застосування обох розроблених топологій наведено у роботі [20].

Відмінністю щодо структури обох модифікацій є різна кількість нейронів в прихованому шарі. У першому випадку вона рівна кількості обраних користувачем центрів RBF. Така топологія нейронної мережі передбачає повну заміну початкових входів задачі на нові, сформовані RBF-функціями. Ця модифікація демонструє високу точність роботи порівняно із базовим лінійним варіантом МПГП, проте характеризується відсутністю екстраполятивних властивостей.

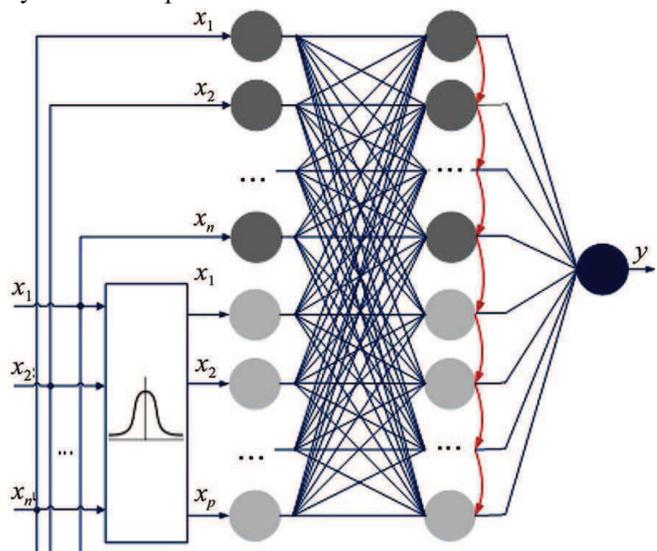


Рис. 1. Топологія комбінованої нейроподібної структури МПГП з додатковим RBF шаром

Друга модифікація (рис. 1) передбачає використання як RBF-входів так і початкових (лінійних) входів. Тому кількість нейронів прихованого шару буде більшою. Це забезпечує ряд переваг такої топології порівняно з першою. Зокрема, за рахунок лінійних входів відбувається відтворення лінійної складової задачі, а входи RBF мо-

делюють нелінійний залишок. Внаслідок цього ця топологія володіє екстраполятивними властивостями. Саме тому її обрано як базовий елемент для розроблення сте-кінгового ансамблю.

Стекінговий ансамбль на підставі однієї нейронної мережі. У цій роботі розроблено стекінговий ансамблевий метод для підвищення точності розв'язання задачі прогнозування. Він базується на використанні комбінованої нейроподібної структури МПГП з RBF розширенням входів. Ансамбль може містити довільну кількість однакових нейромереж цього типу. Вхідні дані для кожного члена ансамблю також будуть однаковими. Змінимо буде тільки кортеж гіперпараметрів комбінованої нейроподібної структури МПГП з RBF розширенням входів, до якого входять: центри RBF; центри мас.

Зміна центрів RBF забезпечує можливість отримання високої точності моделювання у випадку використання як однієї нейронної мережі так і ансамблю [15]. Додаткове застосування різних варіантів зміщення центрів мас до якогось випадково обраного вектора забезпечує можливість підвищення точності прогнозу. Сумісне використання двох цих параметрів забезпечує необхідне різноманіття членів нового ансамблю, кожен з яких формуватиме відмінний від інших вихідний сигнал із використанням одних і тих же навчальних даних.

Структурну схему роботи розробленого ансамблю наведено на рис. 2. Остаточний результат роботи ансамблю формується як середнє арифметичне значення вихідних сигналів усіх його членів.

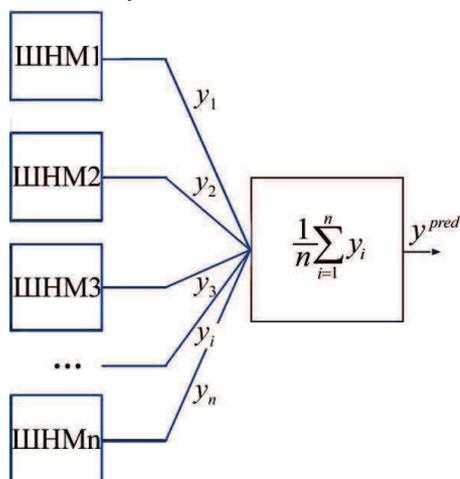


Рис. 2. Структурна схема роботи ансамблю на підставі комбінованої нейроподібної структури МПГП з RBF розширенням входів

Результати дослідження та їх обговорення

1. Опис, аналіз та підготовка набору даних. Для моделювання роботи методу авторами обрано реальний набір даних із сайту Kaggle [10]. Набір містить 1338 (n) спостережень про величину медичних страхових виплат для осіб, що проживають у США. Кожен із векторів містить 6 вхідних атрибутів та 1 вихідний (табл. 1). Задача полягає у прогнозуванні індивідуальних медичних витрат на страхування здоров'я.

Розглянемо детальніше усі незалежні змінні та їх характеристики [19]:

- Insurance contractor age (Age). Мінімальний вік особи яка застраховується становить 18 років, максимальний 64 роки, середній вік по усій вибірці становить 39,2 роки. Цей стовпець містить 574 молодих застрахованих осіб (18-35), 548 осіб середнього віку (35-55) і 216 осіб старшого віку (>55);
- Insurance contractor gender (Gender). Стать застрахованої особи. Набір даних містить 662 жінок (F) та 676 чоловіків (M);
- Body mass index (BMI), kg/m^2 . Це співвідношення висоти людини до її ваги. Мінімальний BMI становить 15,96, максимальний 53,12 та середній 30,66 що є вище норми;
- Number of dependents (Children). Це кількість дітей, які покриваються страховим полісом. Цей показник коливається від 1 до 5, а середнє значення становить 2;
- Smoking column (Smoker). Набір даних містить 1064 курців та 274 не курців;
- Beneficiary's residential area in the US column (Area). У цьому стовпці представлено чотири регіони США, де кількість спостережень по кожному з них становить: північний схід – 324, північний захід – 325, південний схід – 364, південний захід – 325.

Незалежною змінною виступає individual insurance costs (ІС). Для розв'язання поставленої задачі здійснено ряд перетворень вхідних даних з метою представлення їх з текстової у числову форму. Зокрема, додано 5 нових стовпців так: кожен із стовпців Insurance contractor gender та Smoking перетворено у два – male (M) та female (F), smoker та no-smoker відповідно. Стовпець Beneficiary's residential area in the US перетворено у 4 різні, кожен з яких відповідатиме за одну із чотирьох регіонів США: Area 1 південний захід, Area 2 південний схід, Area 3 північний захід і Area 4 північний схід. В такий спосіб отримано нову вибірку даних, кожен із 1338 векторів якої містить 11 вхідних числових атрибутів (табл. 2). Експериментальне моделювання роботи розробленого ансамблю відбувалося із використанням набору даних і табл. 2. Набір даних було випадково розподілено на навчальну та тестову вибірки у співвідношенні 80 до 20 %.

Табл. 1. Частина початкового набору даних [19]

| № | Insurance contractor age | Insurance contractor gender | Body mass index, kg/m^2 | Number of dependents | Smoking | Beneficiary's residential area in the US | Individual insurance costs |
|-----|--------------------------|-----------------------------|---|----------------------|---------|--|----------------------------|
| 1 | 19 | female | 27,9 | 0 | yes | southwest | 16884,924 |
| 2 | 18 | male | 33,77 | 1 | no | southeast | 1725,5523 |
| 3 | 28 | male | 33 | 3 | no | southeast | 4449,462 |
| 4 | 33 | male | 22,705 | 0 | no | northwest | 21984,4706 |
| 5 | 32 | male | 28,88 | 0 | no | northwest | 3866,8552 |
| 6 | 31 | female | 25,74 | 0 | no | northeast | 3756,6216 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| i | 60 | female | 27,9 | 0 | yes | southwest | 16884,924 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| n | 61 | female | 29,07 | 0 | yes | northwest | 29141.3603 |

Табл. 2. Частина перетвореного набору даних [19]

| № | Age | F | M | BMI ₁ , kg/m ² | Child-rens | Smoker | Non-smoker | Area 1 | Area 2 | Area 3 | Area 4 | ПС |
|----------|-----|-----|-----|--------------------------------------|------------|--------|------------|--------|--------|--------|--------|-----------|
| 1 | 19 | 1 | 0 | 27,9 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 16884,924 |
| 2 | 18 | 0 | 1 | 33,77 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1725,5523 |
| 3 | 28 | 0 | 1 | 33 | 3 | 0 | 1 | 0 | 1 | 0 | 0 | 4449,462 |
| 4 | 33 | 0 | 1 | 22,705 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 21984,471 |
| 5 | 32 | 0 | 1 | 28,88 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 3866,8552 |
| 6 | 31 | 1 | 0 | 25,74 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 3756,6216 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| <i>i</i> | 60 | 1 | 0 | 27,9 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 16884,924 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| <i>n</i> | 61 | 0 | 0 | 29,07 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 29141,360 |

Основними параметрами комп'ютера, на якому відбувалися експерименти були: Memory: 8 Gb Intel® Core(TM) i5-6200U CPU, 230GHz, 2.40 GHz.

2. Підбір оптимальний параметрів роботи ансамблю. Стекінговий ансамбль побудовано на використанні однієї і тієї ж навчальної вибірки кожним членом ансамблю. Окрім цього, як члени ансамблю використовують один тип неітеративної нейронної мережі. Кожен член ансамблю відрізняється тільки двома гіперпараметрами, значення яких обираються випадково. Тому основним параметром роботи ансамблю який варто задавати користувачу є кількість членів ансамблю, які забезпечуватимуть найвищу точність роботи.

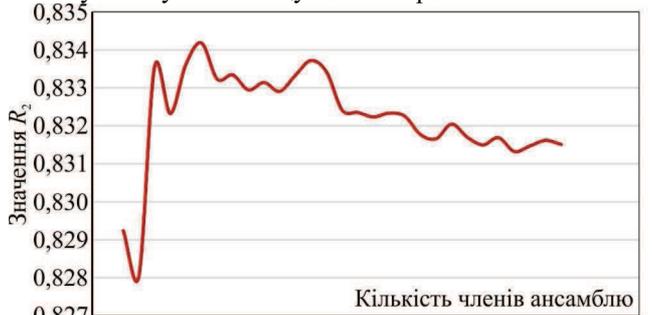


Рис. 3. Зміна значень коефіцієнта детермінації при зміні кількості членів ансамблю

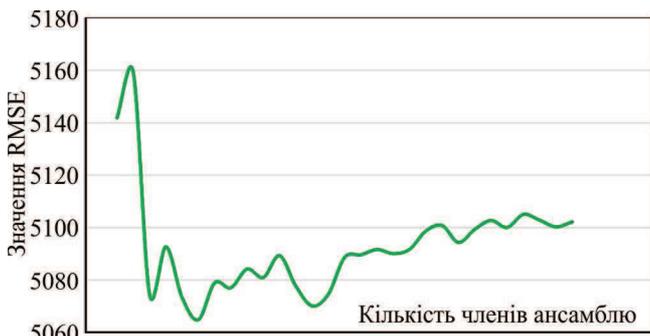


Рис. 4. Зміна значень середньоквадратичної похибки (RMSE) при зміні кількості членів ансамблю

У роботі проведено експериментальне дослідження щодо точності роботи розробленого ансамблю при зміні кількості його членів з 2 до 30. Кількість центрів RBF кожного члена ансамблю рівна 100. Результати такого експерименту при визначенні найбільшого значення коефіцієнту детермінації та найменшого значення середньоквадратичної похибки наведено на рис. 3 та 4.

Як видно з обох рисунків, під час опрацювання обраного набору даних, найвищу точність отримано при виборі семи членів ансамблю.

3. Результати моделювання. Числові результати роботи ансамблю при визначеному оптимальному значенні кількості його елементів зведено у табл. 3.

Табл. 3. Результати моделювання розробленого ансамблю при оптимальних параметрах його роботи

| Кількість членів ансамблю | Коефіцієнт детермінації (R2) | Середньоквадратична похибка (RMSE) | Середня абсолютна похибка у відсотках (MAPE) |
|---------------------------|------------------------------|------------------------------------|--|
| 7 | 0,834182 | 5064,864 | 29,77738 |

Саме їх буде використано під час порівняння ефективності роботи розробленого ансамблю з іншими методами цього класу.

Обговорення результатів дослідження. Результати роботи за розробленим методом порівнювалися із результатами роботи за наявними методами. Ряд методів цього класу, які обрано для порівняння наведено в табл. 4. Результати порівняння на підставі похибок RMSE як в режимах навчання так і застосування наведено на рис. 5. Як видно з рис. 5 наявний бустинговий ансамбль AdaBoost демонструє один з найменш точних результатів. Окрім цього тут спостерігається проблема перенавчання. Це ж характерно і для методу стохастичного градієнтного спуску та багаторівневого перцептрону.

Неітеративні методи машинного навчання демонструють дещо вищу точність роботи. Зокрема комітети нейроподібних структур з додатковим RBF шаром, які побудовано на підставі дихотомії набору даних демонструють високу точність прогнозу. Проте найвищу точність прогнозу медичних страхових виплат отримано саме із використанням розробленого ансамблю.

Окрім цього, у роботі порівнювалася тривалість процедури навчання усіх досліджуваних методів. Результати цього дослідження наведено на рис. 6. Варто зазначити, що тривалість роботи ансамблю визначалася найбільшою величиною тривалості навчання одного із усіх його членів. Такий підхід пояснюється можливістю розпаралелювання, де тривалість процедури навчання усього ансамблю можна визначити саме так.

Як видно з рис. 6, найменшу тривалість процедури навчання демонструє метод стохастичного градієнтного спуску. Проте, точність роботи цього методу дуже низька. Незважаючи на неітеративність процедури навчання, найбільша тривалість навчання показують комітети нейроподібних структур з додатковим RBF шаром. Це пояснюється особливостями їх алгоритмічної реалізації та багаторівневої структури. Розроблений ансамбль демонструє задовільні часові показники для реалізації процедури навчання серед інших аналогічних методів цього класу.

Подальший розвиток цього підходу можливий в чітких межах, поданих у роботі [22], [23], [24].
напряму його апаратної реалізації на підставі напра-

Табл. 4. Обрані для порівняння наявні методи прогнозування

| № п/п | Назва методу | Посилання |
|-------|--|-----------|
| 1 | Стохастичний градієнтний спуск | [7] |
| 2 | Алгоритм AdaBoost | [4] |
| 3 | Багат шаровий перцептрон | [14] |
| 4 | Нейроподібна структура МПГП | [18] |
| 5 | Нейроподібна структура МПГП з додатковим RBF-шаром | [20] |
| 6 | Комбінована нейроподібна структура МПГП з додатковим RBF-шаром | [20] |
| 7 | Комітет на підставі нейроподібної структури МПГП з додатковим RBF-шаром | [8] |
| 8 | Комітет на підставі комбінованої нейроподібної структури МПГП з додатковим RBF-шаром | [21] |

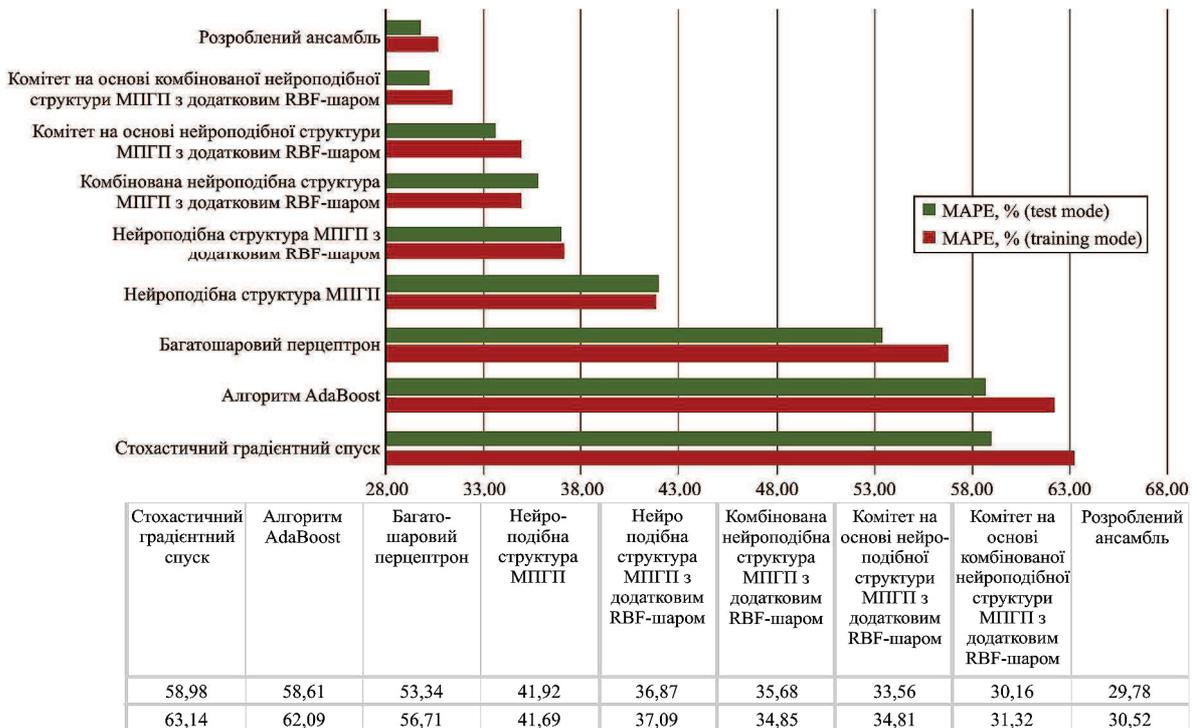


Рис. 5. Значення середньої абсолютної похибки у відсотках (MAPE) для режимів навчання і застосування усіх досліджуваних методів

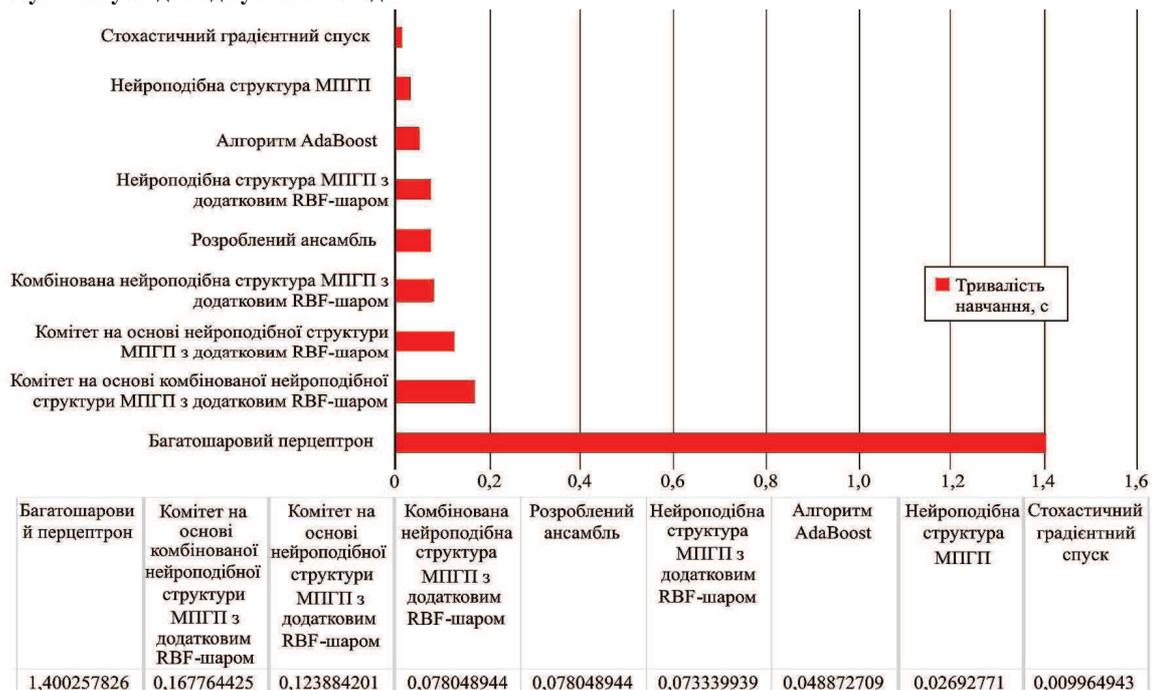


Рис. 6. Тривалість процедури навчання (у секундах) усіх досліджуваних методів

Висновки

Описано просту та ефективну схему ансамблювання нелінійних штучних нейронних мереж неітеративного типу. Авторами розроблено стекінговий ансамбль комбінованої нейроподібної структури МППП з додатковим RBF шаром для підвищення точності розв'язання завдань прогнозування. Розроблений ансамбль передбачає використання елементів одного типу, що навчаються на одному і тому ж наборі даних. Різноманіття елементів ансамблю забезпечено за рахунок використання кожним з них окремого випадкового кортежу гіперпараметрів. Моделювання роботи методу проведено із використанням реального набору даних. Задача полягає у прогнозуванні величини медичних страхових виплат. Встановлено оптимальні параметри роботи ансамблю. Визначено найвищі показники точності розробленого ансамблю порівняно з іншими методами цього класу.

References

- [1] Agarwal, S., & Chowdary, C. R. (2020). A-Stacking and A-Bagging: Adaptive versions of ensemble learning algorithms for spoof fingerprint detection. *Expert Systems with Applications*, 146, 113160. <https://doi.org/10.1016/j.eswa.2019.113160>
- [2] Boodhun, N., & Jayabalan, M. (2018). Risk prediction in life insurance industry using supervised learning algorithms. *Complex & Intelligent Systems*, 4(2), 145–154. <https://doi.org/10.1007/s40747-018-0072-1>
- [3] Chaurasia, V., & Pal, S. (2021). Stacking-Based Ensemble Framework and Feature Selection Technique for the Detection of Breast Cancer. *SN Computer Science*, 2(2), 67. <https://doi.org/10.1007/s42979-021-00465-3>
- [4] Feng, D.-C., Liu, Z.-T., Wang, X.-D., Chen, Y., Chang, J.-Q., Wei, D.-F., & Jiang, Z.-M. (2020). Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach. *Construction and Building Materials*, 230, 117000. <https://doi.org/10.1016/j.conbuildmat.2019.117000>
- [5] Folberth, C., Elliott, J., Müller, C., Balkovič, J., Chryssanthacopoulos, J., Izaurralde, R. C., Jones, C. D., Khabarov, N., Liu, W., Reddy, A., Schmid, E., Skalský, R., Yang, H., Arneth, A., Ciais, P., Deryng, D., Lawrence, P. J., Olin, S., Pugh, T. A. M., ... Wang, X. (2019). Parameterization-induced uncertainties and impacts of crop management harmonization in a global gridded crop model ensemble. *PLOS ONE*, 14(9), e0221862. <https://doi.org/10.1371/journal.pone.0221862>
- [6] Hassan, A. H. A., & Elfaki, E. (2018). Prediction of Electrical Output Power of Combined Cycle Power Plant Using Regression ANN Model. <https://doi.org/10.5281/zenodo.1285164>
- [7] Ighalo, J. O., Adeniyi, A. G., & Marques, G. (2020). Application of linear regression algorithm and stochastic gradient descent in a machine – learning environment for predicting biomass higher heating value. *Biofuels, Bioproducts and Biorefining*, 14(6), 1286–1295. <https://doi.org/10.1002/bbb.2140>
- [8] Izonin, I., Tkachenko, R., Kryvinska, N., Gregus, M., Tkachenko, P., & Vitynskyi, P. (2019). Committee of SGTM Neural-Like Structures with RBF kernel for Insurance Cost Prediction Task. 2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON), 1037–1040. <https://doi.org/10.1109/UKRCON.2019.8879905>
- [9] Kurz, C. F., Maier, W., & Rink, C. (2020). A greedy stacking algorithm for model ensembling and domain weighting. *BMC Research Notes*, 13(1), 1–6. <https://doi.org/10.1186/s13104-020-4931-7>
- [10] Medical Cost Personal Datasets. (n.d.). Retrieved 9 December 2018, from <https://www.kaggle.com/mirichoi0218/insurance>
- [11] Pavlyshenko, B. (2018). Using Stacking Approaches for Machine Learning Models. 2018 IEEE Second International Conference on Data Stream Mining Processing (DSMP), 255–258. <https://doi.org/10.1109/DSMP.2018.8478522>
- [12] Pham, K., Kim, D., Park, S., & Choi, H. (2021). Ensemble learning-based classification models for slope stability analysis. *Catena*, 196, 104886. <https://doi.org/10.1016/j.catena.2020.104886>
- [13] Rocca, J. (2021, March 21). *Ensemble methods: Bagging, boosting and stacking*. Medium. <https://towardsdatascience.com/ensemble-methods-bagging-boosting-and-stacking-c9214a10a205>
- [14] Salah, M., Altalla, K., Salah, A., & Abu-Naser, S. S. (2018). Predicting Medical Expenses Using Artificial Neural Network. *International Journal of Engineering and Information Systems (IJEAIS)*, 2(10), 7.
- [15] Shaikhina, T., & Khovanova, N. A. (2017). Handling limited datasets with neural networks in medical applications: A small-data approach. *Artificial Intelligence in Medicine*, 75, 51–63. <https://doi.org/10.1016/j.artmed.2016.12.003>
- [16] Shakhovska, N., Yakovyna, V., & Kryvinska, N. (2020). An Improved Software Defect Prediction Algorithm Using Self-organizing Maps Combined with Hierarchical Clustering and Data Preprocessing. In S. Hartmann, J. Küng, G. Kotsis, A.M. Tjoa, & I. Khalil (Eds.), *Database and Expert Systems Applications* (pp. 414–424). Springer International Publishing. https://doi.org/10.1007/978-3-030-59003-1_27
- [17] Teslyuk, V., Kazarian, A., Kryvinska, N., & Tsmots, I. (2021). Optimal Artificial Neural Network Type Selection Method for Usage in Smart House Systems. *Sensors*, 21(1), 47. <https://doi.org/10.3390/s21010047>
- [18] Tkachenko, R., & Izonin, I. (2019). Model and Principles for the Implementation of Neural-Like Structures Based on Geometric Data Transformations. In Z. Hu, S. Petoukhov, I. Dychka, & M. He (Eds.), *Advances in Computer Science for Engineering and Education*. Vol. 754, (pp. 578–587). Springer International Publishing. https://doi.org/10.1007/978-3-319-91008-6_58
- [19] Tkachenko, R., Izonin, I., Vitynskyi, P., Lotoshynska, N., & Pavlyuk, O. (2018). Development of the Non-Iterative Supervised Learning Predictor Based on the Ito Decomposition and SGTM Neural-Like Structure for Managing Medical Insurance Costs. *Data*, 3(4), 46. <https://doi.org/10.3390/data3040046>
- [20] Tkachenko, R., Kutucu, H., Izonin, I., Doroshenko, A., & Tsymbal, Y. (n.d.). *Non-Iterative Neural-Like Predictor for Solar Energy in Libya*. 11.
- [21] Tkachenko, R., Tkachenko, P., Izonin, I., Vitynskyi, P., Kryvinska, N., & Tsymbal, Y. (2019). Committee of the Combined RBF-SGTM Neural-Like Structures for Prediction Tasks. In I. Awan, M. Younas, P. Ünal, & M. Aleksy (Eds.), *Mobile Web and Intelligent Information Systems* (pp. 267–277). Springer International Publishing. https://doi.org/10.1007/978-3-030-27192-3_21
- [22] Tsmots, I., & Skorokhoda, O. (2010). Methods and VLSI-structures for neural element implementation. 2010 *Proceedings of Vth International Conference on Perspective Technologies and Methods in MEMS Design*, 135–135.
- [23] Tsmots, I., Skorokhoda, O., & Rabyk, V. (2016). *Structure and Software Model of a Parallel-Vertical Multi-Input Adder for FPGA Implementation*, 158–160. <https://doi.org/10.1109/STC-CSIT.2016.7589894>
- [24] Tsmots, I., Teslyuk, V., & Vavruk, I. (2013). Hardware and software tools for motion control of mobile robotic system. 2013 12th International Conference on the Experience of Designing and Application of CAD Systems in Microelectronics (CADSM), 368–368.

[25] Xiao, Y., Wu, J., Lin, Z., & Zhao, X. (2018). A deep learning-based multi-model ensemble method for cancer

prediction. *Computer Methods and Programs in Biomedicine*, 153, 1–9. <https://doi.org/10.1016/j.cmpb.2017.09.005>

R. O. Tkachenko, I. V. Izonin, V. M. Danylyk, V. Yu. Mykhalevych

Lviv Polytechnic National University, Lviv, Ukraine

STACKING OF THE SGTМ NEURAL-LIKE STRUCTURE WITH RBF LAYER BASED ON GENERATION OF A RANDOM CURTAIN OF ITS HYPERPARAMETERS FOR PREDICTION TASKS

Improving prediction accuracy by artificial intelligence tools is an important task in various industries, economics, medicine. Ensemble learning is one of the possible options to solve this task. In particular, the construction of stacking models based on different machine learning methods, or using different parts of the existing data set demonstrates high prediction accuracy of the. However, the need for proper selection of ensemble members, their optimal parameters, etc., necessitates large time costs for the construction of such models. This paper proposes a slightly different approach to building a simple but effective ensemble method. The authors developed a new model of stacking of nonlinear SGTМ neural-like structures, which is based on the use of only one type of ANN as an element base of the ensemble and the use of the same training sample for all members of the ensemble. This approach provides a number of advantages over the procedures for building ensembles based on different machine learning methods, at least in the direction of selecting the optimal parameters for each of them. In our case, a tuple of random hyperparameters for each individual member of the ensemble was used as the basis of ensemble. That is, the training of each combined SGTМ neural-like structure with an additional RBF layer, as a separate member of the ensemble occurs using different, randomly selected values of RBF centers and centers of mass. This provides the necessary variety of ensemble elements. Experimental studies on the effectiveness of the developed ensemble were conducted using a real data set. The task is to predict the amount of health insurance costs based on a number of independent attributes. The optimal number of ensemble members is determined experimentally, which provides the highest prediction accuracy. The results of the work of the developed ensemble are compared with the existing methods of this class. The highest prediction accuracy of the developed ensemble at satisfactory duration of procedure of its training is established.

Keywords: stacking; ensemble learning; neural-like structures; radial-basis functions; medical insurance costs.

Інформація про авторів:

Ткаченко Роман Олексійович, д-р техн. наук, професор, завідувач кафедри інформаційних технологій видавничої справи.

Email: roman.tkachenko@gmail.com; <https://orcid.org/0000-0002-9802-6799>; ResearcherID: R-2695-2017

Ізонін Іван Вікторович, канд. техн. наук, доцент, кафедра систем штучного інтелекту. **Email:** ivanizonin@gmail.com;

<https://orcid.org/0000-0002-9761-0096>; ResearcherID: Q-4678-2016

Данилик Віталій Миколайович, студент, кафедра інформаційних систем і мереж. **Email:** webdvitaly@gmail.com

Михалевич Вікторія Юрїївна, студентка, кафедра систем штучного інтелекту. **Email:** mihalevichvika@gmail.com;

<https://orcid.org/0000-0002-4177-7217>

Цитування за ДСТУ: Ткаченко Р. О., Ізонін І. В., Данилик В. М., Михалевич В. Ю. Стекінг нейроподібної структури МПГП з RBF шаром на підставі генерування випадкового кортежу її гіперпараметрів для завдань прогнозування. *Український журнал інформаційних технологій*. 2021, т. 3, № 1. С. 49–55.

Citation APA: Tkachenko, R. O., Izonin, I. V., Danylyk, V. M., & Mykhalevych, V. Yu. (2021). Stacking of the SGTМ neural-like structure with RBF layer based on generation of a random curtain of its hyperparameters for prediction tasks. *Ukrainian Journal of Information Technology*, 3(1), 49–55. <https://doi.org/10.23939/ujit2021.03.049>