

DOI: 10.20535/1810-0546.2018.2.129022

УДК 004.8

О.С. Мішук*, П.Б. Вітинський

Національний університет “Львівська політехніка”, Львів, Україна

НЕЙРОННА МЕРЕЖА З КОМБІНОВАНОЮ АПРОКСИМАЦІЄЮ ПОВЕРХНІ ВІДГУКУ

Проблематика. Існує велика кількість нейронних мереж, які мають свої переваги та недоліки. Наприклад, прості, швидкі та зручні у використанні одношарові перцептрони придатні лише для лінійних і лінеаризованих задач регресії, а складніші нейронні мережі затратні за часом навчання та прогнозування. Тому виникає завдання розробки швидких та ефективних алгоритмів навчання штучних нейронних мереж. Додатковим фактором для дослідження нових методів навчання нейронних мереж є пошук найменших похибок навчання та прогнозування.

Мета дослідження. Пошук і аналіз властивостей найефективнішого методу навчання штучних нейронних мереж із застосуванням комбінованої апроксимації поверхні відгуку. Виконання обчислювальних експериментів стосовно запропонованих штучних нейронних мереж та порівняння результатів експериментів із відомими та розробленими методами.

Методика реалізації. Поставлена мета досягається завдяки аналізу відомих методів комбінованої апроксимації поверхні відгуку, розробці нових алгоритмів для навчання нейронних мереж на основі кластеризації даних методом k -means та вибору того алгоритму, за якого похибки навчання штучної нейронної мережі та прогнозування даних будуть найменшими.

Результати дослідження. Наведено результати досліджень різних методів навчання штучних нейронних мереж. Проаналізовано особливості методів комбінованої апроксимації поверхні відгуку. Показано, що розроблені два методи комбінованої апроксимації поверхні відгуку для навчання та застосування штучних нейронних мереж підтверджують ефективність запропонованого підходу. Вибрано алгоритм комбінованої апроксимації, який забезпечує найменші похибки навчання та прогнозування.

Висновки. Розроблені методи комбінованої апроксимації поверхні відгуку дають змогу навчати нейронні мережі та прогнозувати дані з меншою похибкою, ніж під час використання моделі авторегресії з ковзним середнім, багатшарового перцептрона чи штучних нейронних мереж моделей геометричних перетворень без додаткової обробки даних.

Ключові слова: нейронна мережа; модель геометричних перетворень; комбінована апроксимація; метод k -means; кластеризація.

Вступ

Нейронні мережі (НМ) успішно застосовуються в різних сферах – бізнесі, медицині, техніці, геології, фізиці тощо, для чого будують машини, які навчаються вирішувати складні завдання, наближаючи й іноді перевершуючи продуктивність на рівні людини [1]. НМ вирішують завдання прогнозування, класифікації, розпізнавання, керування, а підвищення їх ефективності пов'язане із загальними тенденціями покращення характеристик технічних засобів обробки інформації [2].

Штучні нейронні мережі (ШНМ) як засоби інформаційного моделювання забезпечують функцію відображення векторів вхідних сигналів у вектори вихідних, де закон відображення задається поверхнею відгуку, яка формується в процесі навчання ШНМ [3]. Найчастіше використовують ШНМ, що формують лінійні поверхні відгуку (одношарові перцептрони, ме-

режі Adaline і Madaline), східчасті (ШНМ зустрічного поширення), нелінійні поверхні сигмоїдних схилів (багатшарові перцептрони), нелінійні поверхні відгуку, побудовані на гіперсферах (нейромережі радіальних базових функцій) [4].

Кожен із перелічених типів ШНМ має специфічні властивості в плані складності настроювання, часу навчання, співвідношення характеристик точність–швидкодія. В плані надійності та швидкості навчання, що має особливе значення для задач великих даних, найбільш привабливими є одношарові перцептрони, мережі Adaline і Madaline, однак їх застосування обмежується класом лінійно розділених задач. ШНМ зустрічного поширення також швидко й ефективно навчаються, в тому числі для задач із нелінійними поверхнями відгуку, але похибки застосування для них у більшості випадків значно перевищують допустимі значення.

*corresponding author: oleksandra.myroniuk@gmail.com

Отже, виникає задача дослідження ШНМ із комбінованою апроксимацією поверхні відгуку для знаходження алгоритму, що максимально задовольнятиме потреби у збільшенні швидкості, ефективності й точності навчання НМ.

ШНМ із комбінованою апроксимацією поверхні відгуку присвячені роботи таких науковців, як А.А. Севастьянов, С.С. Харинцев, М.Х. Салахов та ін. [3], які досліджували алгоритми комбінованої апроксимації у визначених сферах використання, наприклад у задачах прикладної спектроскопії. Але підбір алгоритмів комбінованої апроксимації залежить від особливостей даних, тому для кожної сфери використання необхідно проводити свої експерименти для покращення результатів навчання і прогнозування.

Прості та зручні у використанні НМ (ОШП, Adaline), які дають змогу оперувати з великими даними, на жаль, придатні лише для лінійних і лінеаризованих задач, а НМ інших типів мають обмеження щодо розмірності завдань і вимагають достатньо високої кваліфікації користувачів. Отже, актуальними є побудова архітектури та розробка ефективних і швидких алгоритмів навчання ШНМ.

Постановка задачі

Мета роботи – вдосконалити структури і методи навчання ШНМ із застосуванням комбінованої апроксимації поверхні відгуку; дослідити їх властивості для задач регресії, а саме

виконати обчислювальні експерименти стосовно запропонованих ШНМ та порівняти результати експериментів із відомими моделями.

Теоретичні основи методу навчання

НМ увійшли в практику скрізь, де потрібно вирішувати завдання прогнозування, класифікації, розпізнавання, керування. Такий успіх визначається насамперед великими можливостями. По-перше, НМ – потужний метод моделювання, що дає змогу відтворювати надзвичайно складні залежності [5]. Крім того, НМ справляються із “прокляттям розмірності”, яке не дає можливості моделювати лінійні залежності в разі великого числа змінних [6]. По-друге, НМ є простими у використанні, адже створюються на прикладах. Користувач НМ підбирає представницькі дані, а потім запускає алгоритм навчання, який автоматично сприймає структуру даних. Від користувача, очевидно, потрібен певний набір евристичних знань про те, як слід відбирати та готувати дані, вибирати потрібну архітектуру мережі й інтерпретувати результати [2].

У дослідженні виконуємо реальну задачу прогнозування споживання електроенергії з навчальною (рис. 1, а) і тренувальною (рис. 1, б) вибірками даних. Навчальна вибірка складається з 365 векторів, що включають у себе отримані дані за тотожну кількість днів попереднього року. Тренувальна вибірка складається з 214 векторів, що описують відповідну кількість днів наступного року, для яких

x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	y
0,992160248	0,854418964	1	0	10	0	0	0	11,4	84	0,969444444	266067,7054
0,992922595	0,851830192	0	0	30	0	0	0	14,8	88	0,858333333	399213,7541
0,993650134	0,849284178	0	0	20	0	0	0	13,7	94	1,344444444	308883,6862
0,994342017	0,846781407	0	0	10	0	0	0	16,2	98	0,288888889	246431,2379
0,994997417	0,844322355	0	0	50	0	1	0	14,4	84	0,352777778	455089,7725
0,995615537	0,841907489	0	0	60	0	0	0	11,1	97	0,305555556	469969,8012

а

x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	y
0,991363963	0,854418964	0	0	40	0	0	1	10,8	91	1,313888889	464373,0739
0,992160248	0,851830192	0	0	20	0	0	0	11,4	99	0,84	316854,1401
0,992922595	0,849284178	1	0	10	0	0	0	7,9	87	0,561111111	239680,2697
0,993650134	0,846781407	1	0	10	0	0	0	8,5	83	0,519444444	242339,7119
0,994342017	0,844322355	0	0	50	0	1	0	10,5	89	0,497222222	455924,5373
0,994997417	0,841907489	0	0	60	0	0	0	12,8	86	0,916666667	464680,6157

б

Рис. 1. Вибірка даних досліджуваної задачі: а – еталонна (навчальна); б – тестова

виконувалося прогнозування в режимі тестування. Вектори обох вибірок даних містять у собі 11 вхідних ознак x_{ij} (показники стану електричної мережі, отримані за даними телеметрії) та 1 вихід y_i (щоденні значення спожитої електроенергії).

Для виконання поставленої задачі прогнозування вибираємо використання штучної нейронної мережі моделі геометричних перетворень (ШНМ МГП) лінійного типу, особливість якої полягає в тому, що вибрана НМ передбачає функціонування лише в лінійному режимі, де здійснюється апроксимація поверхні відгуку у вигляді площини (гіперплощини). Лінійна ШНМ МГП забезпечує дещо вищу точність і одночасно більшу швидкість навчання порівняно з багатошаровим перцептроном.

Опис методів. У дослідженні використовуємо відому НМ “ANN-RBF”, де можна вибрати кілька варіантів налаштування і отримати результати прогнозування у вигляді графіка та похибок MAPE, RMSE і RMSE_M. За допомогою експериментів із налаштуванням різних режимів роботи мережі було визначено, що використання режиму No RBF є найдоцільнішим, тому що при виконанні навчання та прогнозування НМ у режимі RBF1 інколи результуює з меншими похибками, ніж в інших режимах, але важливим негативним фактором є серйозні часові затримки. При використанні режиму RBF2 часові затримки найменші, але тут похибки навчання та прогнозування є дуже

значними. При виконанні ж програми у режимі No RBF похибки прогнозування та часові затримки найменші. Наприклад, загальна похибка MAPE для навчання становить 3,88 %, а похибка прогнозування – 4,19 %, що можна побачити на рис. 2.

У багатьох випадках прогнозування використовуються методи лінійної апроксимації, оскільки більшість реальних процесів у бізнесі та економіці характеризуються саме лінійними моделями. Але не завжди точність наближення лише лінійними методами є задовільною, тому для підвищення точності використовуємо комбінації різних методів.

У дослідженні також виконуємо кластеризацію вхідних векторів наборів даних. Існує багато алгоритмів кластеризації, серед яких можна виділити ієрархічний алгоритм, k -means, c -means, виділення зв'язних компонент, поширену кластеризацію. Для нашого дослідження вибрали найоптимальніший алгоритм k -means, оскільки в ньому можна задати потрібну кількість кластерів.

Алгоритм k -means розділяє вхідну множину на кластери так, щоб середні значення в кластерах були максимально різними. Середні значення називаються центрами, або прототипами, кластерів [7]. Для запуску алгоритму необхідно знати число k – передбачуване число кластерів. Часто алгоритм запускають на одних і тих же даних багато разів із різними значеннями k , а потім вибирають найбільш відповідний результат [8].

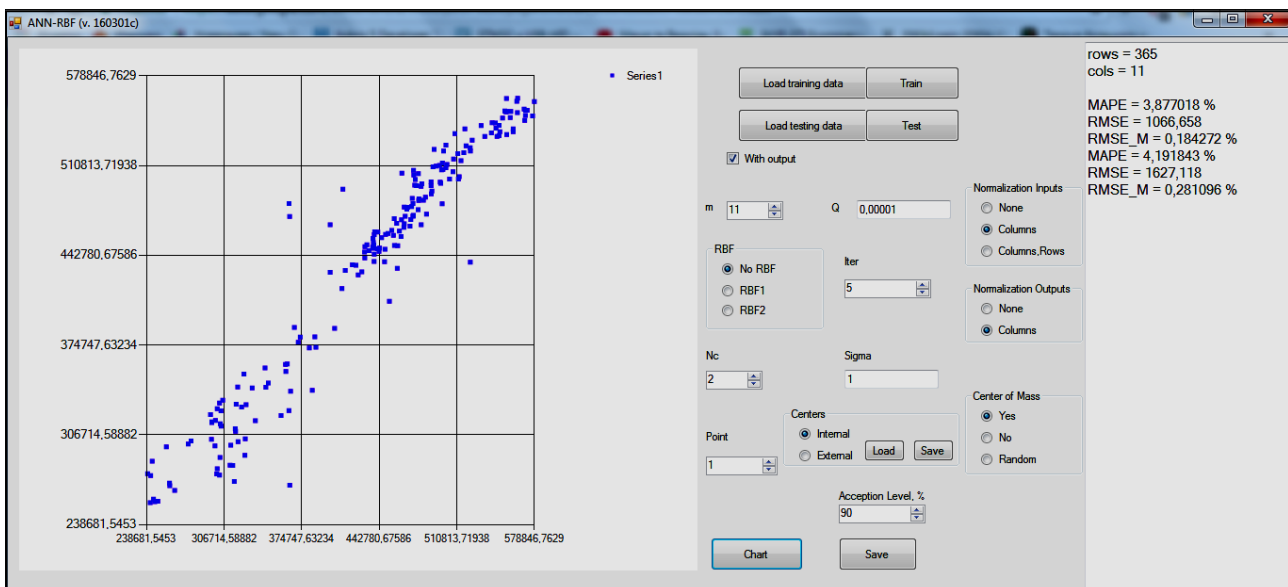


Рис. 2. Початковий результат навчання та прогнозування в режимі “No RBF”

Алгоритм кластеризації k -means має таку послідовність:

1. Розподіляємо точки по кластерах випадковим чином.

2. Випадковим чином вибираємо k спостережень, які на цьому кроці вважаємо центрами кластерів.

3. Кожну з точок множини відносимо до того кластера, відстань до центра якого для неї мінімальна.

4. Якщо на попередньому кроці жодна точка не перейшла в інший кластер або якщо ми зробили максимально допустиме число ітерацій, припиняємо роботу. Інакше переходимо до пункту 2.

5. Результатом роботи алгоритму є поточна точка розбиття на кластери [9].

Першим пунктом, використовуваним у комбінованій апроксимації, виконуємо нормування даних (вхідних векторів x_i , що показано на рис. 3): знаходимо максимальне значення в кожному стовпці та присвоюємо йому одиницю, решту значень ділимо на знайдене максимальне значення за модулем.

Оскільки підбір алгоритмів для використання в комбінованій апроксимації залежить від особливостей даних, то в дослідженні найоптимальніших комбінацій методів для кращого результату і точнішого прогнозування були виконані експерименти з кількома варіантами кластеризації.

Розглянемо перший алгоритм комбінованої апроксимації (метод 1).

1. Кластеризуємо методом k -means навчальну матрицю по 11 еталонних входах на k компактних множин точок, де k – кількість кластерів.

2. Знаходимо центри мас кожного кластера за формулами

$$c_x = \frac{\sum_i m_i x_i}{\sum_i m_i},$$

$$c_y = \frac{\sum_i m_i y_i}{\sum_i m_i},$$

де c_x та c_y – координати точки вектора центрів мас.

3. Знаходимо Евклідову відстань d за формулою

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2},$$

де x_i – точка з вектора тестової (тренувальної) матриці, а y_i – точка з вектора центрів мас кластера з навчальної вибірки.

4. Кожен вектор із тестової вибірки даних відносимо до того кластера, Евклідова відстань до вектора центрів мас якого є найменшою. Таким чином кластеризуємо всі вектори тренувальної матриці.

5. Виконуємо розширення матриць, до входів кожного вектора навчальної та тренувальної вибірок додаємо k додаткових входів, де k – вибрана кількість кластерів. Якщо вектор належить до k -го кластера, то k -му додатковому входу буде присвоєно 1, а решті нулі, тому отримуємо вибірку векторів (рис. 4) вигляду $x_{i1}, \dots, x_{i11}, 0, \dots, 1, \dots, 0, \dots, 0$.

6. Розширені навчальну та тренувальну вибірки даних, де вектори мають $11 + k$ входів і 1 вихід, подаємо на ШНМ "ANN-RBF" та знаходимо похибки MAPE (Middle Absolute Percent Error) – середні абсолютні похибки у відсотках для навчання та прогнозування.

Описаний алгоритм виконуємо визначену кількість разів, щоб знайти похибки MAPE (табл. 1) для різної заданої кількості кластерів.

Другий досліджуваний алгоритм комбінованої апроксимації (метод 2) складається з таких кроків:

1. Маючи пронормовані навчальну та тренувальну вибірки векторів вигляду x_{i1}, \dots, x_{i11} ,

x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	y
0,992160248	0,679754853	1	0	0,166666667	0	0	0	0,511210762	0,84	0,622103387	266067,7054
0,992922595	0,677695289	0	0	0,5	0	0	0	0,66367713	0,88	0,550802139	399213,7541
0,993650134	0,675669743	0	0	0,333333333	0	0	0	0,614349776	0,94	0,862745098	308883,6862
0,994342017	0,673678599	0	0	0,166666667	0	0	0	0,726457399	0,98	0,185383244	246431,2379
0,994997417	0,671722238	0	0	0,833333333	0	1	0	0,64573991	0,84	0,226381462	455089,7725
0,995615537	0,669801029	0	0	1	0	0	0	0,497757848	0,97	0,196078432	469969,8012
0,996195608	0,667915337	0	0	1	0	0	0	0,506726457	0,94	0,290552585	473983,3085
0,996736899	0,666065518	0	0	1	0	0	0	0,381165919	0,99	0,932263815	472153,0707

Рис. 3. Нормована навчальна вибірка даних

x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	k1	k2	k3	k4	k5	y
0,99216	0,67975	1	0	0,16667	0	0	0	0,51121	0,84	0,62210	0	0	1	0	0	266067,7054
0,99292	0,67769	0	0	0,5	0	0	0	0,66367	0,88	0,55080	0	0	1	0	0	399213,7541
0,99365	0,67566	0	0	0,33333	0	0	0	0,61434	0,94	0,86274	0	0	1	0	0	308883,6862
0,99434	0,67367	0	0	0,16667	0	0	0	0,72645	0,98	0,18538	0	0	1	0	0	246431,2379
0,99499	0,67172	0	0	0,83333	0	1	0	0,64573	0,84	0,22638	1	0	0	0	0	455089,7725
0,99561	0,66980	0	0	1	0	0	0	0,49775	0,97	0,19607	0	0	1	0	0	469969,8012
0,99619	0,66791	0	0	1	0	0	0	0,50672	0,94	0,29055	0	0	1	0	0	473983,3085
0,99673	0,66606	0	0	1	0	0	0	0,38116	0,99	0,93226	0	0	1	0	0	472153,0707
0,99723	0,66425	0	0	0,66667	0	0	1	0,71748	0,67	0,88413	0	0	0	0	1	464338,8536

Рис. 4. Розширена навчальна вибірка векторів

Таблиця 1. Порівняння результатів виконання досліджених методів

Кількість кластерів	Метод 1		Метод 2	
	Похибка при навчанні, %	Похибка при прогнозуванні, %	Похибка при навчанні, %	Похибка при прогнозуванні, %
k = 5	3,770	4,045	3,744	3,878
k = 10	3,882	4,075	3,685	3,736
k = 15	3,438	3,570	3,275	3,304
k = 20	3,399	3,525	3,107	3,120
k = 25	3,064	3,215	3,072	3,152

y_i , формуємо об'єднану вибірку векторів (365 + 214) вигляду x_{i1}, \dots, x_{i11} без урахування виходів.

2. Кластеризуємо методом *k-means* об'єднану матрицю, яка складається з 579 векторів, на *k* кластерів.

3. Виконуємо розділення об'єднаної вибірки назад на навчальну та тренувальну матриці з відомими відношеннями векторів до певних кластерів.

4. Розширюємо матриці аналогічно попередньому алгоритму: до входів кожного вектора досліджуваних вибірок додаємо *k* додаткових входів, де *k*-му додатковому входу буде присвоєно 1, якщо вектор належить до *k*-го кластера, а решті входів присвоюємо нулі.

5. Навчаємо ШНМ "ANN-RBF" за допомогою розширених вибірки даних, де вектори мають 11 + *k* входів і 1 вихід, та застосовуємо, знаходячи похибки MAPE.

Порівняємо результати розроблених методів. Похибки при виконанні розробленого методу 2 є меншими, ніж похибки, знайдені при виконанні розробленого методу 1, що можна побачити з табл. 1.

Беручи до уваги принципи лінійної суперпозиції, подаємо еквівалентну схему вибраної ШНМ. Вихідний сигнал ШНМ описується формулою

$$y = \sum_{i=1}^n x_i w_i + \sum_{j=1}^L c_j w_j,$$

де, враховуючи спосіб формування вхідних векторів, для кожного вектора лише на одному із входів нейромережі маємо сигнал одиниці й нулі на всіх інших входах. Це означає, що розроблений метод забезпечує комбіновану апроксимацію поверхні відгуку – лінійну з однаковими коефіцієнтами лінійної форми у всій області визначення й одночасно індивідуальні постійні зміщення для кожного виділеного кластера даних.

Порівняємо отримані результати з іншими методами, а саме порівняємо розроблені методи 1 та 2 з моделлю авторегресії з ковзним середнім (АРКС), де похибка прогнозування є найбільшою; багатошаровим перцептроном (БП) та ШНМ МГП (табл. 2).

Таблиця 2. Порівняння результатів виконання з іншими методами

Похибка	АРКС	БП	ШНМ МГП	Метод 1	Метод 2
MAPE	10,9 %	4,2 %	4,19 %	3,69 %	3,44 %

Висновки

За відсутності яскраво виражених коливань вихідного сигналу можна віддати перевагу використанню лінійної мережі для задачі прогнозування, оскільки при цьому забезпечується згладжування даних від випадкової складової.

З виконаного дослідження можна зробити висновок, що застосування ШНМ МГП дає прийнятний рівень апроксимації вихідних спостережуваних даних. Єдиним параметром настроювання є задавання кількості кластерів. Досліджено, що зі збільшенням кількості кластерів похибка спочатку зменшується, а потім починає рости.

ШНМ із комбінованою апроксимацією забезпечують належне узагальнення, отже, ви-

бираючи кількість кластерів, які забезпечують мінімальну похибку при навчанні, отримуємо невелику похибку при прогнозуванні. Наприклад, при класифікації по 20 кластерах похибка при навчанні становить 3,107 %, а похибка при прогнозуванні – 3,12 %, що на 1 % менше похибки при прогнозуванні без апроксимації поверхні відгуку. В деяких випадках похибка прогнозування може бути меншою за похибку навчання.

Подальші дослідження будуть спрямовані на аналіз нових алгоритмів комбінованої апроксимації даних для знаходження мінімальної похибки навчання та прогнозування.

References

- [1] M.I. Jordan and T.M. Mitchell, "Machine learning: trends, perspectives, and prospects", *Science*, vol. 349, no. 6245, pp. 255–260, 2015. doi: 10.1126/science.aaa8415
- [2] R. Tkachenko, *Neural Network Means of Artificial Intelligence*. Lviv, Ukraine: Lviv Polytechnic Publ., 2017.
- [3] A. Sevastjanov *et al.* (2003). Neural Network Regularization of the Solution of Inverse Ill-Posed Applications of Applied Spectroscopy [Online]. Available: <http://zhurnal.ape.relarn.ru/articles/2003/189.pdf>
- [4] R. Tkachenko, "New paradigm of artificial neural networks of direct distribution", *Visnyk Natsionalnoho Universytetu "Lvivska politekhnika": Komp'uterna Inzheneriia ta Informatsiini Tekhnologii*, no. 386, pp. 43–54, 1999.
- [5] R. Kruse *et al.*, *Computational Intelligence: A Methodological Introduction*. London: Springer, 2013.
- [6] C. Zhang *et al.*, "Understanding deep learning requires rethinking generalization", in *Proc. 5th Int. Conf. Learning Representations (ICLR)*, 2017.
- [7] U.V. Volosiuk, "Analysis of clustering algorithms for data analysis tasks", *Proc. of the Military Institute of Kyiv National Taras Shevchenko University*, no. 47, pp. 112–119, 2014.
- [8] O.I. Derkach, "Analytical processing of text information with the help of clustering tools", *Molodyi Vchenyi: Physyko-Matematychni Nauky*, no. 7, pp. 159–165, 2016.
- [9] L. Jiang *et al.*, "Survey of improving k-nearest-neighbor for classification", in *Proc. 4th Int. Conf. FSKD 2007*, vol. 1, pp. 679–683, 2007, doi: 10.1109/FSKD.2007.552.

О.С. Мищук, П.Б. Витинский

НЕЙРОННАЯ СЕТЬ С КОМБИНИРОВАННОЙ АППРОКСИМАЦИЕЙ ПОВЕРХНОСТИ ОТКЛИКА

Проблематика. Существует большое количество нейронных сетей, которые имеют свои преимущества и недостатки. Например, простые, быстрые и удобные в использовании однослойные перцептроны пригодны только для линейных и линеаризованных задач регрессии, а более сложные нейронные сети расходны по времени обучения и прогнозирования. Поэтому возникает задача разработки быстрых и эффективных алгоритмов обучения искусственных нейронных сетей. Дополнительным фактором для исследования новых методов обучения нейронных сетей является поиск наименьших ошибок обучения и прогнозирования.

Цель исследования. Поиск и анализ свойств максимально эффективного метода обучения искусственных нейронных сетей с применением комбинированной апроксимации поверхности отклика. Выполнение вычислительных экспериментов относительно предложенных искусственных нейронных сетей и сравнение результатов экспериментов с известными и разработанными методами.

Методика реализации. Поставленная цель достигается благодаря анализу известных методов комбинированной апроксимации поверхности отклика, разработке новых алгоритмов для обучения нейронных сетей на основе кластеризации данных методом *k-means* и выбору того алгоритма, при котором погрешности обучения искусственной нейронной сети и прогнозирования данных будут наименьшими.

Результаты исследования. Приведены результаты исследований различных методов обучения искусственных нейронных сетей. Проанализированы особенности методов комбинированной апроксимации поверхности отклика. Показано, что разработанные два метода комбинированной апроксимации поверхности отклика для обучения и применения искусственных нейронных сетей подтверждают эффективность предложенного подхода. Выбран алгоритм комбинированной апроксимации, который обеспечивает наименьшие погрешности обучения и прогнозирования.

Выводы. Разработанные методы комбинированной аппроксимации поверхности отклика позволяют обучать нейронные сети и прогнозировать данные с меньшей погрешностью, чем при использовании модели авторегрессии со скользящим средним, многослойного перцептрона или искусственных нейронных сетей моделей геометрических преобразований без дополнительной обработки данных.

Ключевые слова: нейронная сеть; модель геометрических преобразований; комбинированная аппроксимация; метод *k*-means, кластеризация.

O.S. Mischuk, P.B. Vitynski

NEURAL NETWORK WITH COMBINED APPROXIMATION OF THE SURFACE OF THE RESPONSE

Background. There are a large number of neural networks that have their advantages and disadvantages, for example, simple, fast and easy to use single-stranded perceptrons are suitable for linear and linearized regression tasks, and more complicated neural networks are expendable in training and prediction time. Therefore, the problem arises for the development of fast and efficient algorithms for training artificial neural networks. An additional factor for researching new methods for training neural networks is finding the smallest training and prediction errors.

Objective. The aim of the paper is to search and analyze the properties of the most effective method of training artificial neural networks using a combined approximation of the response surface. Another step is to perform computational experiments on proposed artificial neural networks and compare the results of experiments with known and developed methods.

Methods. Analysis of known methods of combined approximation of the response surface was used. New algorithms for training neural networks, based on clustering of data using *k*-means method were developed. The algorithm with the smallest errors of artificial neural network learning and data prediction is chosen.

Results. The results of research of different methods of training of artificial neural networks are given. Peculiarities of the methods of combined approximation of the response surface are analyzed. It is shown that the two methods of combined approximation of the response surface for training of artificial neural networks and prediction confirm the effectiveness of the proposed approach. Combined approximation algorithm is selected, which provides the lowest learning and forecasting errors.

Conclusions. It was investigated that developed methods of combined approximation of the response surface allow training neural networks and predicting data with less error than when using autoregressive model with moving average, multilayer perceptron or artificial neural networks of models of geometric transformations without additional data processing.

Keywords: neural network; model of geometric transformations; combined approximation; *k*-means method; clusterization.

Рекомендована Радою
Навчально-наукового комплексу
“Інститут прикладного системного
аналізу” КПІ ім. Ігоря Сікорського

Надійшла до редакції
21 лютого 2018 року

Прийнята до публікації
29 березня 2018 року