

УДК 004.9

О. В. Корнієнко¹, С. О. Субботін¹, О. Е. Наривський²

¹Запорізький національний технічний університет,
вул. Жуковського, 64, 69063 Запоріжжя, Україна

²ТОВ «Укрспецмаш»

вул. Гагаріна, 7, 71111 Бердянськ, Україна

Нейромережеве моделювання критичних температур підтінгоутворення сталі

Вирішено завдання створення математичного забезпечення для побудови моделей кількісних залежностей на основі нейронних мереж прямого поширення. Запропоновано метод виключення нейронів, що враховує вплив кожного нейрона на помилку моделі. Вирішено практичне завдання визначення критичних температур підтінгоутворення сталі AiSi 321 за її характеристиками. Виконано побудову нейромережевих моделей, їхнє навчання та тестування на даних за характеристиками сталі. Порівняно результати тестування всіх побудованих моделей.

Ключові слова: *вибірка, навчання, нейронна мережа, помилка, градієнт, виключення, ймовірність, сталь AiSi 321.*

Вступ

Теплообмінне обладнання знаходить широке використання у сучасній енергетиці, хімічній і нафтогазовій промисловості.

У виробництві теплообмінного обладнання часто використовують корозійностійкі сталі та сплави, леговані хромом, нікелем і молібденом. Ці конструкційні матеріали мають високу корозійну стійкість у багатьох агресивних середовищах, але можуть піддаватися підтінговій корозії в оборотних водах, що містять хлорид, які застосовують для охолодження технологічних продуктів у теплообмінному обладнанні [1].

Тому актуальною проблемою є побудова моделей залежності критичних температур підтінгоутворення (КТП) від характеристик використовуваної сталі і оборотної води [1].

Перспективним засобом побудови таких моделей є штучні нейронні мережі. Завдяки їхнім здатностям до навчання за прецедентами, узагальнення даних і витягу знань з даних нейронні мережі можуть використовуватися для побудови моделей залежностей за кількісними спостереженнями в умовах відсутніх або невідомих аналітичних залежностей [2].

© О. В. Корнієнко, С. О. Субботін, О. Е. Наривський

Серед відомих архітектур нейронних мереж найбільш потужною і одночасно доступною для подальшого аналізу та сприйняття людиною є багат шарові нейронні мережі прямого поширення сигналу, які обрано як базис у даній роботі.

Проте недоліками моделей, які побудовані на основі нейронних мереж прямого поширення сигналу, є їхня надлишковість, що також негативно впливає на інтерпретабельність побудованих моделей.

Отже, актуальним завданням є удосконалення методів побудови моделей на основі нейромереж прямого поширення сигналу.

Метою роботи було удосконалення математичного забезпечення побудови нейромоделей та їхнє застосування для вирішення задачі моделювання критичних температур п'іттінгоутворення сталі.

Постановка задачі

Нехай ми маємо вибірку спостережень $x = \langle x^s, y^s \rangle$, де x^s — s -й екземпляр вибірки; $x^s = \{x_j^s\}$, x_j^s — значення j -ї вхідної ознаки, що характеризує s -й екземпляр вибірки; $j = 1, 2, \dots, N$, $s = 1, 2, \dots, S_{all}$, де N — кількість вхідних ознак, що характеризують екземпляри вибірки; S_{all} — кількість екземплярів у вибірці.

Тоді задача побудови моделі залежності $y = f(w, x)$ на основі нейронної мережі із заданою користувачем структурою f буде полягати в знаходженні оптимальних значень набору вагових коефіцієнтів нейронів мережі w , тобто таких, при яких критерій якості моделі — середньоквадратична помилка E буде мінімальною.

Середньоквадратична помилка нейромоделі E визначається за формулою [3]

$$E = \sqrt{\sum_{i=1}^Y (R_i - M_i)^2}, \quad (1)$$

де R_i — реальне значення вихідного параметра y_i^s ; M_i — розрахункове значення параметра y_i^s , що отримане за допомогою моделі; $i = 1, \dots, Y$ (Y — кількість вихідних параметрів моделі).

Архітектура багат шарової нейромережі

Нейронна мережа прямого поширення має шарувату архітектуру (рис. 1) [4]: у таких моделях вхідні ознаки розпізнаваного екземпляра подаються на входи мережі та далі поширюються від вхідного шару (x_1, \dots, x_n) до вихідного (y_1, \dots, y_m) .

Кожен нейрон на прихованих і вихідному шарах отримує на входах сигнали з виходів нейронів попереднього шару. Вхідний вектор кожного нейрона перетворюється у скаляр за допомогою вагової (дискримінантної) функції, після чого значення вихідного сигналу розраховується за допомогою функції активації [5]. Вихідні значення останнього шару є результатом роботи нейромережі.

Вихідні значення кожного нейрона визначаються за формулою:

$$O_m = \psi(w_{m,0} + \sum_{n=1}^N w_{m,n} \cdot I_n), \quad (2)$$

де $\psi(x)$ — функція активації; w — вагові коефіцієнти; I — вихідні значення нейронів попереднього шару; N — кількість нейронів попереднього шару, $m = 1, 2, \dots, M$ (M — кількість нейронів поточного шару).

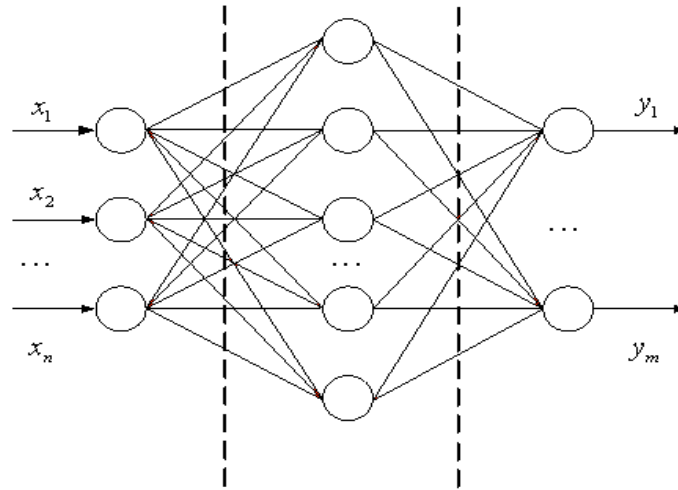


Рис. 1. Багатошаровий перцептрон

Навчання нейромережі із контрастуванням і видаленням нейронів

Навчання нейронних мереж прямого поширення, як правило, проводиться на основі градієнтного методу [6] та техніки зворотного поширення помилки [7]. У процесі навчання мінімізується функція помилки мережі E , яка визначає різницю між виходами мережі та цільовими значеннями. Як функцію помилки доцільно використовувати середньоквадратичну помилку E , що обчислюється за формулою (1).

Вектор R визначається залежно від класу екземпляра і функції активації нейронів вихідного шару нейромережі. Значення R_i , що відповідає класу екземпляра, повинне бути близьким до максимального значення функції активації, а всі інші — близькими до мінімального. Якщо область значень функції активації необмежена, доцільно використовувати значення 1 та -1 або 1 та 0. У задачах регресії значення R_i відповідають цільовим значенням вихідних параметрів.

Метод навчання нейронної мережі полягає у послідовному виконанні таких етапів.

Етап ініціалізації. Задати навчальну вибірку даних $\langle x, y \rangle$ та структуру нейромережі. Ваговим коефіцієнтам нейронів задати випадкові значення з діапазону $[-0,1; 0,1]$. Задати: кількість циклів (епох) навчання $Epochs = 0$, максимальну кількість циклів (епох) навчання $MaxEpochs$, крок навчання α [8].

Етап корекції ваг. Послідовно у циклі для $Epochs = 0, 1, \dots, MaxEpochs$, де $MaxEpochs$ — максимально допустима задана кількість епох навчання, виконуєть-

ся навчання мережі на основі екземплярів вибірки шляхом виконання етапу навчання мережі на основі екземпляра та коректується крок навчання α .

Етап навчання мережі на основі екземплярів вибірки. Послідовно на входи нейромережі подається i -й екземпляр навчальної вибірки, $i = 1, 2, \dots, S_{tr}$, де S_{tr} — кількість екземплярів навчальної вибірки. Для поточного екземпляра розраховуються вихідні значення нейронної мережі за формулою (2). В процесі розрахування важливо зберегти вихідні значення нейронів усіх шарів. Після цього обчислюється середньоквадратична помилка мережі E .

Далі обчислюються градієнти помилок нейронів вихідного шару:

$$G_i = (O_i - R_i) \cdot (1 - O_i^2), \quad i = 0, \dots, I,$$

де I — кількість вихідних нейронів; O — вихідний вектор нейронної мережі; R — бажаний вихідний вектор; G — градієнти помилок вихідного шару.

Якщо як функцію активації ψ використовувати функцію гіперболічний тангенс [4]

$$\text{Tanh}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}},$$

то її похідна після спрощення матиме вигляд $1 - \text{Tanh}(x)^2$, а в даному випадку O — це вже обчислене значення $\text{Tanh}(x)$.

Після цього обчислюються градієнти для нейронів прихованих шарів:

$$G_i = \sum_{j=0}^J Gr_j \cdot w_{i,j} \cdot (1 - O_i^2), \quad i = 0, \dots, I, j = 0, \dots, J,$$

де I — кількість нейронів поточного шару; J — кількість нейронів наступного шару; Gr — градієнти помилок наступного шару; G — градієнти помилок поточного шару; O — вихідний вектор поточного шару; w — вагові коефіцієнти нейронів наступного шару.

Після розрахунку градієнтів необхідно виконати корекцію вагових коефіцієнтів за формулою

$$w_{i,j} = w_{i,j} - O_i \cdot G_j \cdot \alpha, \quad i = 0, \dots, I, j = 0, \dots, J,$$

де I — кількість нейронів поточного шару; J — кількість нейронів наступного шару; G — градієнти помилок; O — вихідний вектор поточного шару; w — вагові коефіцієнти нейронів поточного шару.

Етап контрастування нейронної мережі. Контрастування полягає у видаленні частини найменш значимих зв'язків у нейронній мережі. Видалення слабких зв'язків приведе до збільшення узагальнюючих властивостей нейромережі [9]. Для видалення зв'язку достатньо для відповідного вагового коефіцієнта задати нульове значення. Отже, для проведення контрастування нейронної мережі необхідно для всіх вагових коефіцієнтів $|w_{j,k}^i| < P$, де $i = 1, \dots, I - 1, j = 1, \dots, J, k = 1, \dots, K$, задати нульове значення (I — кількість шарів нейронної мережі, J — кількість нейронів i шару; K — кількість нейронів $i + 1$ шару; P — порогове значення для вагових коефіцієнтів).

Одним із методів оптимізації процесу навчання нейронних мереж є виключення нейронів [10]. Він полягає у тому, що в процесі навчання деяка частина нейронів видаляється з мережі. На кожній ітерації навчання виключаються випадкові нейрони. Кожен нейрон може бути виключеним з імовірністю P . Імовірність виключення може бути різною для кожного шару нейронної мережі. На рис. 2 зображено звичайну нейронну мережу, а на рис. 3 — нейронну мережу після виключення нейронів.

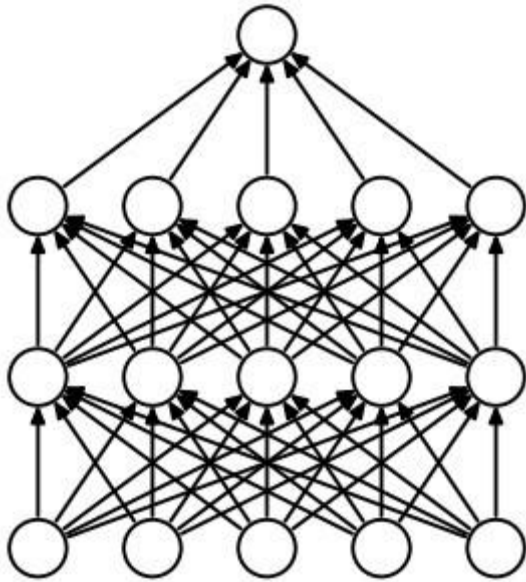


Рис. 2. Звичайна нейронна мережа

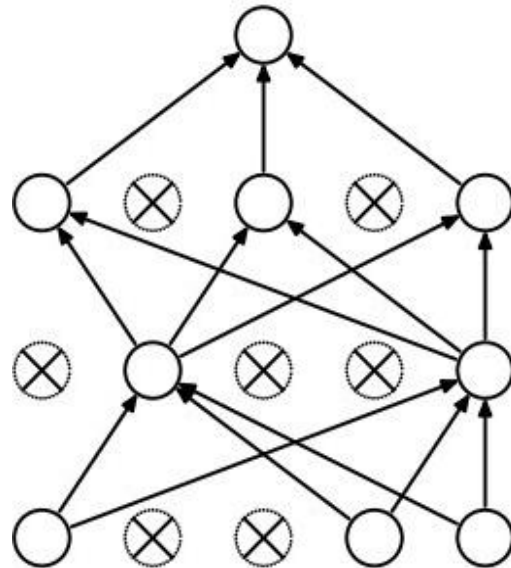


Рис. 3. Нейронна мережа після виключення нейронів

Виключення нейронів проводиться тільки при навчанні моделі. При реалізації методу достатньо прирівняти до 0 вихідні значення тих нейронів, що повинні бути виключені на ітерації у процесі навчання.

При тестуванні та під час роботи моделі необхідно використовувати всі нейрони. При цьому вихідні значення нейронів на шарах, де використовувалося виключення, будуть визначатися за формулою [10]

$$O_m = Q\psi(w_{m,0} + \sum_{n=1}^N w_{m,n} \cdot I_n),$$

де $Q = 1 - P$ (P — ймовірність виключення нейронів на шарі); $\psi(x)$ — функція активації; w — вагові коефіцієнти; I — вихідні значення нейронів попереднього шару; N — кількість нейронів попереднього шару; $m = 1, 2, \dots, M$ (M — кількість нейронів) поточного шару.

Виключення нейронів дозволяє уникнути перенавчання моделі [11], що в більшості випадків приводить до збільшення точності моделі на тестових даних.

Для збільшення точності моделі пропонується визначати ймовірність виключення нейронів за їхнім впливом на помилку моделі:

$$P_n = \frac{P_b}{2} + \frac{P_b}{1 + e^{\frac{2Ck}{N} - k}},$$

де P_b — базова ймовірність виключення нейронів на шарі; N — кількість нейронів на шарі; C — кількість нейронів, для яких абсолютні значення градієнтів більші, ніж для n -го нейрона; k — коефіцієнт, що впливає на нелінійність залежності. Рекомендується обирати значення k з інтервалу $[1 \dots 10]$.

Таким чином, імовірність виключення нейронів зростає зі збільшенням їхнього впливу на помилку.

При навчанні нейромережі з модифікованим методом виключення нейронів необхідно двічі розраховувати вихідні значення нейронів і градієнти. В перший раз вихідні значення та градієнти розраховуються із використанням усіх нейронів. Після цього визначаються нейрони, що повинні бути виключеними на ітерації. Далі проводяться етапи розрахування вихідних значень, розрахування градієнтів та етап корекції вагових коефіцієнтів з урахуванням виключених нейронів. При тестуванні та під час роботи нейромережі коефіцієнт Q , на який помножується функція активації, визначається за формулою $Q = 1 - P_b$, де P_b — базова ймовірність виключення нейронів.

Експерименти та результати

Для побудови нейромережових моделей було розроблено програмне забезпечення мовою C#.

Програмне забезпечення дозволяє виконувати побудову нейронних мереж, їхнє навчання та тестування на вибірках даних.

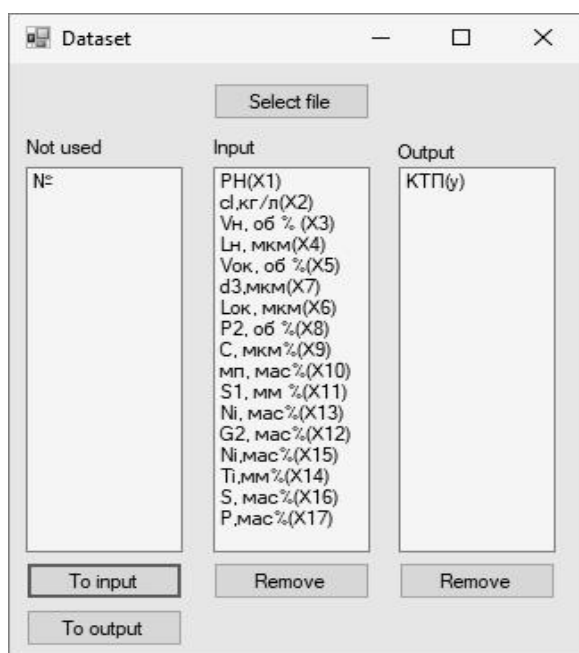


Рис. 4. Вибір вхідних і вихідних параметрів

Для початку навчання нейромережі необхідно завантажити вибірку з файлу у форматі csv, після чого обрати вхідні та вихідні параметри (рис. 4). В першому рядку файлу повинні міститися назви параметрів.

Після вибору даних необхідно обрати частини вибірки для навчання та тестування, тип задачі, структуру нейромережі та інші параметри (рис. 5).

У процесі навчання будується графік залежності середньоквадратичної помилки на навчальних даних від епохи.

Після завершення навчання виконується тестування, після чого створюється текстовий файл з детальною інформацією про результати тестування.

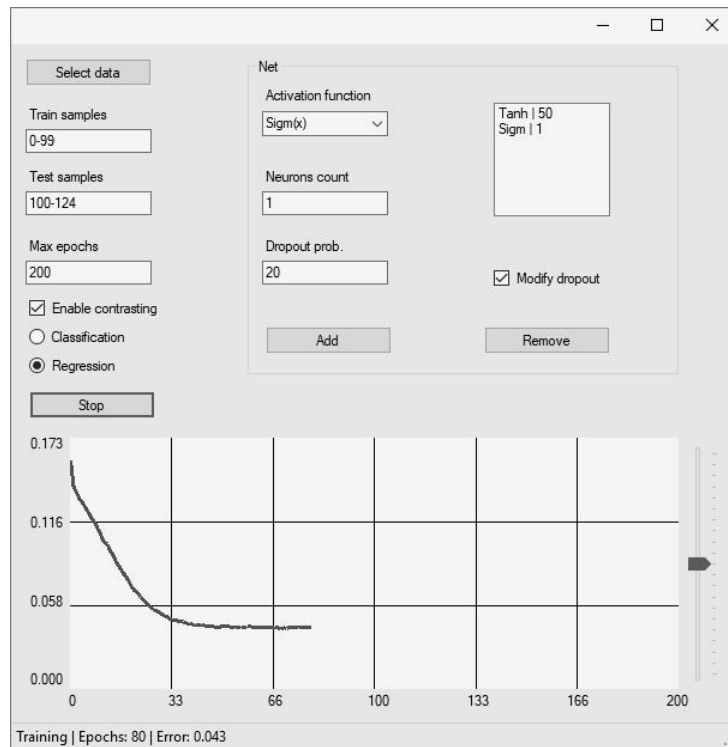


Рис. 5. Вибір параметрів навчання

Вирішення задачі побудови моделі сталі AiSi 321 від характеристик сталі та оборотної води здійснювалося за даними експериментальних спостережень. Вибірка даних складалася зі 125 екземплярів — точкових спостережень, що характеризувалися значеннями вхідних ознак: x_1 — рН оборотної води; x_2 — вміст хлоридів в оборотній воді (Cl^-); x_3 — обсяг нітридів титану (V_H); x_4 — середня відстань між нітридами титану (L_H); x_5 — обсяг оксидів (V_{OK}); x_6 — середня відстань між оксидами (L_{OK}); x_7 — середній діаметр зерна аустеніту (d_3); x_8 — об'єм δ -фериту (P_α); x_9 — вміст вуглецю (C) у сталі; x_{10} — вміст марганцю (Mn) у сталі; x_{11} — вміст кремнію (Si) у сталі; x_{12} — вміст хрому (Cr) в сталі; x_{13} — вміст нікелю (Ni) у сталі; x_{14} — вміст титану (Ti) у сталі; x_{15} — вміст азоту (N) у сталі; x_{16} — вміст сірки (S) у сталі; x_{17} — вміст фосфору (P) у сталі; а як вихідна ознака у використовувався КТП.

Вибірку даних було нормовано шляхом відображення значень ознак на інтервал $[0; 1]$. Нові значення визначалися за формулою:

$$x_i^s = \frac{x_i^s - x_i^{\min}}{x_i^{\max} - x_i^{\min}},$$

де x_i^{\max} , x_i^{\min} — максимальне та мінімальне значення i -ї вхідної ознаки серед усіх екземплярів вибірки, відповідно. Аналогічно визначалися нормовані значення для вихідного параметра y .

Для проведення експериментів нормовану вибірку було розділено на навчальну та тестову частини шляхом вибору 25 випадкових екземплярів для тестової частини. Тестову частину нормованої вибірки наведено в табл. 1.

Таблиця 1. Фрагмент вибірки даних

КТП	pH	Cl	V _H	L _H	V _{OK}	L _{OK}	d3	P2	C	Mn	Si	Cr	Ni	Ti	N	S	P
0,171	0,000	0,800	0,333	0,680	0,818	0,106	0,707	1,000	1,000	0,000	0,440	0,787	1,000	1,000	0,500	0,000	0,000
0,561	0,750	0,200	1,000	0,440	0,227	0,616	1,000	0,042	0,147	0,917	0,520	0,528	0,000	0,100	0,500	0,000	0,000
0,512	0,250	0,000	1,000	0,440	0,227	0,616	1,000	0,042	0,147	0,917	0,520	0,528	0,000	0,100	0,500	0,000	0,000
0,780	1,000	0,200	0,265	0,000	0,242	0,603	0,000	0,165	0,294	1,000	0,320	1,000	0,000	0,400	1,000	0,000	0,000
0,366	0,500	0,600	0,693	1,000	1,000	0,000	0,379	1,000	0,000	0,771	1,000	0,000	0,067	0,300	0,000	1,000	0,500
0,683	1,000	0,000	0,693	1,000	1,000	0,000	0,379	1,000	0,000	0,771	1,000	0,000	0,067	0,300	0,000	1,000	0,500
0,585	0,250	0,000	0,333	0,680	0,818	0,106	0,707	1,000	1,000	0,000	0,440	0,787	1,000	1,000	0,500	0,000	0,000
0,756	1,000	0,200	0,333	0,680	0,818	0,106	0,707	1,000	1,000	0,000	0,440	0,787	1,000	1,000	0,500	0,000	0,000
0,512	0,000	0,200	0,265	0,000	0,242	0,603	0,000	0,165	0,294	1,000	0,320	1,000	0,000	0,400	1,000	0,000	0,000
0,512	1,000	0,600	0,693	1,000	1,000	0,000	0,379	1,000	0,000	0,771	1,000	0,000	0,067	0,300	0,000	1,000	0,500
0,439	0,000	0,000	1,000	0,440	0,227	0,616	1,000	0,042	0,147	0,917	0,520	0,528	0,000	0,100	0,500	0,000	0,000
0,195	0,000	0,600	0,693	1,000	1,000	0,000	0,379	1,000	0,000	0,771	1,000	0,000	0,067	0,300	0,000	1,000	0,500
0,488	0,000	0,000	0,333	0,680	0,818	0,106	0,707	1,000	1,000	0,000	0,440	0,787	1,000	1,000	0,500	0,000	0,000
0,073	0,000	1,000	0,333	0,680	0,818	0,106	0,707	1,000	1,000	0,000	0,440	0,787	1,000	1,000	0,500	0,000	0,000
0,366	0,500	0,800	1,000	0,440	0,227	0,616	1,000	0,042	0,147	0,917	0,520	0,528	0,000	0,100	0,500	0,000	0,000
0,415	0,750	0,600	0,333	0,680	0,818	0,106	0,707	1,000	1,000	0,000	0,440	0,787	1,000	1,000	0,500	0,000	0,000
0,878	0,750	0,000	0,000	0,440	0,000	1,000	0,129	0,000	0,000	0,833	0,000	0,772	0,233	0,000	1,000	1,000	1,000
0,805	0,750	0,000	0,333	0,680	0,818	0,106	0,707	1,000	1,000	0,000	0,440	0,787	1,000	1,000	0,500	0,000	0,000
1,000	1,000	0,000	0,333	0,680	0,818	0,106	0,707	1,000	1,000	0,000	0,440	0,787	1,000	1,000	0,500	0,000	0,000
0,829	0,750	0,000	0,265	0,000	0,242	0,603	0,000	0,165	0,294	1,000	0,320	1,000	0,000	0,400	1,000	0,000	0,000
0,488	0,250	0,200	0,693	1,000	1,000	0,000	0,379	1,000	0,000	0,771	1,000	0,000	0,067	0,300	0,000	1,000	0,500
0,317	0,500	0,600	0,333	0,680	0,818	0,106	0,707	1,000	1,000	0,000	0,440	0,787	1,000	1,000	0,500	0,000	0,000
0,585	0,250	0,200	0,000	0,440	0,000	1,000	0,129	0,000	0,000	0,833	0,000	0,772	0,233	0,000	1,000	1,000	1,000
0,707	0,500	0,000	0,333	0,680	0,818	0,106	0,707	1,000	1,000	0,000	0,440	0,787	1,000	1,000	0,500	0,000	0,000
0,854	1,000	0,000	0,265	0,000	0,242	0,603	0,000	0,165	0,294	1,000	0,320	1,000	0,000	0,400	1,000	0,000	0,000

У табл. 2 наведено результати експериментів. Для всіх експериментів використовувалася тришарова нейромережа. Функція активації прихованого шару — гіперболічний тангенс, вихідного шару — сигмоїдна логістична. Кількість вхідних нейронів дорівнює кількості вхідних параметрів, кількість вихідних нейронів — 1, кількість нейронів прихованого шару змінюється залежно від експерименту.

Табл. 2. Результати експериментів

№ експерименту	Кількість нейронів прихованого шару	Базова ймовірність виключення нейронів	Час навчання (с)	Коефіцієнт узагальнення	Середня помилка
1	20	20 %	9	4,68	0,0481
			9	4,71	0,0460
			16	4,72	0,0427
2	50	20 %	13	1,89	0,0477
			13	1,88	0,0448
			26	1,87	0,0424
3	50	30%	13	1,89	0,0477
			12	1,88	0,0447
			26	1,88	0,0432
4	80	20 %	16	1,16	0,0474
			16	1,18	0,0448
			37	1,17	0,0432
5	80	40 %	16	1,16	0,0474
			15	1,17	0,0454
			37	1,18	0,0452

Обсяг оперативної пам'яті для всіх моделей не перевищував 1 МБ. Для всіх моделей було використано однакові значення кількості епох і кроку навчання.

Для порівняння моделей як один із критеріїв використовувався коефіцієнт узагальнення моделлю навчальних даних $K_{уз}$, що визначається за формулою

$$K_{уз} = \frac{NS}{N_w},$$

де N — кількість ознак; S — кількість екземплярів в навчальній вибірці; N_w — кількість вагових коефіцієнтів з ненульовими значеннями.

Для зменшення впливу випадкових значень вагових коефіцієнтів було зроблено генератор псевдовипадкових чисел [12], що дозволяє використовувати однакові значення вагових коефіцієнтів у межах експерименту. В межах кожного експерименту наведено по 3 значення: перше показує результати для нейромережі без виключення нейронів, друге — для звичайного методу виключення, третє — для модифікованого методу виключення. Як критерій якості використано середнє арифметичне значення середньоквадратичних помилок для кожного екземпляра тестової частини вибірки.

У табл. 3 наведено реальні значення вихідного параметра та значення, отримані за допомогою побудованих нейромережових моделей. Використано моделі з експерименту 2, тому що з ними отримано найкращі результати. Всі дані наведено для тестової частини вибірки.

Таблиця 3. Значення вихідних параметрів

№ екземпляра	Модель 1	Модель 2	Модель 3	Реальне значення
1	0,1418	0,1424	0,1283	0,1707
2	0,6203	0,6190	0,6295	0,5610
3	0,5372	0,5376	0,5432	0,5122
4	0,8024	0,8017	0,8213	0,7805
5	0,3506	0,3511	0,3429	0,3659
6	0,8036	0,7944	0,8104	0,6829
7	0,5474	0,5499	0,5601	0,5854
8	0,7065	0,7100	0,7308	0,7561
9	0,4793	0,4826	0,4850	0,5122
10	0,5352	0,5336	0,5387	0,5122
11	0,4433	0,4442	0,4430	0,4390
12	0,2076	0,2058	0,1895	0,1951
13	0,4553	0,457	0,4595	0,4878
14	0,1045	0,1015	0,0867	0,0732
15	0,2388	0,2459	0,2356	0,3659
16	0,4297	0,4286	0,4272	0,4146
17	0,8179	0,8237	0,8437	0,8780
18	0,7116	0,7179	0,7409	0,8049
19	0,7693	0,7837	0,8117	1,0000
20	0,8057	0,8076	0,8284	0,8293
21	0,4577	0,4594	0,4598	0,4878
22	0,3429	0,3404	0,3321	0,3171
23	0,6010	0,6029	0,6097	0,5854
24	0,6337	0,6388	0,6562	0,7073
25	0,8561	0,8556	0,8753	0,8537

За результатами експериментів можна зробити висновок, що запропонована модифікація методу виключення нейронів, не зважаючи на збільшення часу навчання, дозволяє підвищити точність нейромоделей. При цьому максимальна ефективність методу у порівнянні зі звичайним методом виключення досягається при невеликій ймовірності виключення нейронів.

Висновки

У роботі вирішено актуальне завдання удосконалення математичного забезпечення для побудови нейромоделей та моделювання залежності критичних температур пітінгоутворення сталі AiSi 321 від характеристик сталі та оборотної води.

Наукова новизна роботи полягає у тому, що запропоновано модифікацію методу виключення нейронів у нейромережах прямого поширення. Модифікація методу полягає у визначенні ймовірності виключення кожного нейрона в залежності від його впливу на помилку мережі. Це дозволяє збільшити точність побудованих моделей.

Практична цінність отриманих результатів полягає у тому, що: розроблено програмне забезпечення, яке реалізує запропонований метод і дозволяє виконувати побудову нейронних мереж, їхнє навчання та тестування на вибірках даних; вирішено практичне завдання визначення критичних температур пітінгоутворення сталі AiSi 321 за її характеристиками та характеристиками оборотної води.

Перспективи подальших досліджень полягають у визначенні оптимальних параметрів нейронних мереж і параметрів виключення нейронів, а також у застосуванні запропонованого метода на більш широкому класі задач.

1. Наривский А.Э., Субботин С.А., Беликов С.Б., Яр-Мухамедова Г.Ш., Атчибаев Р.А. Питтингостойкость стали АіSi 321 в хлоридсодержащих средах в зависимости от ее состава и структурной гетерогенности. *ҚазҰТЗУ Хабаршысы (Вестник КазННТУ)*. 2019. 1. С. 240–250.
2. Руденко О.Г., Бодянский С.В. Штучні нейронні мережі. Харків: Компанія СМІТ, 2006. 404 с.
3. Bishop C. M. *Pattern recognition and machine learning*. Springer, 2006. 738 p.
4. Хайкин С. Нейронные сети: полный курс. Москва: Вильямс, 2006. 1104 с.
5. Kumar K.V. *Neural networks and fuzzy logic*. S.K. Kataria & Sons, 2016. 300 p.
6. Bertsekas D. P. *Convex Optimization Theory*. Athena Scientific, 2009. 256 p.
7. Nielsen M. *Neural Networks and Deep Learning*. Determination Press, 2015. 216 p.
8. Goodfellow I., Bengio Y., Courville A. *Deep Learning*. MIT Press, 2016. 781 p.
9. Louizos C., Welling M., Kingma D. P. Learning sparse neural networks through L_0 regularization. *International Conference on Learning Representations*, 2018.
10. Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdinov R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*. 2014. **15**. P. 1929–1958.
11. Caruana R., Lawrence S., Giles L. Overfitting in Neural Nets: Backpropagation, Conjugate Gradient, and Early Stopping. *Neural information processing systems conference*, 2000.
12. L'Ecuyer P. *Random Number Generation*. Springer, 2007. 31 p.

Надійшла до редакції 05.03.2019