

УДК 004.032.26(045)

М.Ю. Буриченко, О.Б. Іванець

Національний авіаційний університет, Київ, Україна

ШЛЯХИ ЗМЕНШЕННЯ НЕВИЗНАЧЕНОСТІ ПРОГНОЗУ СТАНУ ОРГАНІЗМУ ЛЮДИНИ ПРИ НЕЙРОМЕРЕЖОВОМУ МОДЕЛЮВАННІ

В роботі проведено аналіз джерел невизначеності, що виникають при побудові моделей для вирішення завдань прогнозування діагностичних ознак біологічних об'єктів за допомогою штучних нейронних мереж. Запропоновано рекомендації щодо зменшення невизначеності прогнозу на різних етапах побудови нейромереж.

Ключові слова: невизначеність, штучні нейронні мережі, моделі прогнозування, медико-біологічні параметри, діагностичні ознаки.

Вступ

Пророцтва майбутніх результатів рідко виявляються точними, тому при прогнозуванні можлива лише спроба пом'якшити наслідки неминучих помилок. Процес діагностування та прогнозування стану, наприклад, серцево-судинної системи, достатньо складний, має різнохарактерні показники та велику кількість факторів невизначеності: вихідних даних, параметрів зовнішнього середовища, показників, пов'язаних з характером динаміки процесів, що відбуваються в організмі тощо. Тому у медико-біологічній області зазвичай не вдається дати однозначно обґрунтований прогноз. Широке застосування такого інструменту як штучні нейронні мережі (ШНМ) в медицині саме і зумовлено здатністю вирішувати ті проблеми і завдання, які не мають інших алгоритмів для вирішення або доступні рішення занадто складні, щоб їх застосовувати. Можливості ШНМ дозволяють працювати з великим обсягом даних, враховувати більшість факторів впливу на досліджуваний об'єкт і швидко приймати рішення.

Виклад основного матеріалу

У роботі представлено результати побудови ШНМ для прогнозування виникнення мікроінфаркту та часу його виникнення. Для задач прогнозування стану організму в якості вхідних та навчальних даних обиралися клінічні дані стану серцево-судинної системи. В якості тестової вибірки використані дані медичного обстеження 110 пацієнтів кардіологічного відділення Інституту геронтології протягом 2-х років. Ці дані характеризують стан серцево-судинної системи за 17-ма показниками аналізу крові (загалом 1870 параметрів). В якості вихідних

параметрів вказується час від обстеження до виникнення мікроінфаркту t та наявність чи відсутність мікроінфаркту k . Тестова вибірка містить діагностичні дані десяти пацієнтів. При цьому проведено аналіз 128 побудованих ШНМ, виявлено джерела невизначеності, що виникають при побудові ШНМ для прогнозування, причинно-наслідкову діаграму яких наведено на рис. 1.

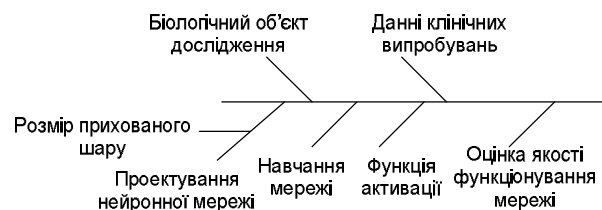


Рис. 1. Діаграма Ісікави

У верхній частині діаграми знаходяться джерела невизначеності, зумовлені вхідними даними. Такими даними є результати клінічних досліджень крові людини, та власне, сам об'єкт дослідження, який є біологічним об'єктом, з притаманними йому нестабільними характеристиками, що обумовлює невизначеність оцінки його стану через зміну впливів внутрішніх та зовнішніх чинників, наприклад, вплив стресу, дія соціального середовища, екологічні катаклізми тощо. Аналіз та особливості оцінювання невизначеності результатів клінічних досліджень розглядається у деяких літературних джерелах, наприклад [1].

Нижня частина діаграми описує джерела невизначеності, що виникають при безпосередній побудові та аналізі вихідних даних ШНМ. Ці джерела умовно можна поділити на 4 групи: некоректне проектування нейронної мережі, оцінка якості функціонування ШНМ, процес навчання мережі та функції активації.

При **проектванні нейронної мережі** необхід-
но в першу чергу вирішити питання про кількість шарів і кількість елементів (нейронів) у кожному шарі. Кількість вхідних і вихідних елементів визначаються властивостями вхідних і вихідних даних, що впливає на визначення розміру прихованого шару. При цьому має місце компроміс між точністю і узагальнюючою спроможністю мережі.

У загальному випадку величина повної похибки, що обумовлює невизначеність при проектуванні нейронної мережі становить:

$$e_i = e_{0i} + \text{reg } e_i, \quad (1)$$

де e_{0i} – похибка, обчислена одним із стандартних способів; $\text{reg } e_i = \lambda/p \cdot \sum_i \omega_i^2$ – функція похибок, що враховує: параметр регуляризації λ , розмір тренувальної послідовності p , ваговий коефіцієнт ω_i .

З виразу (1) видно, чим більше значення набувають λ та ω_i , тим більше зростає похибка. Величина λ представляє з себе відношення сигнал / шум.

Використання мережі за типом «багатошарового перцептрон», як видно з табл. 1, виявляє найменше значення похибки на навчальній та тестовій множині 0,19 та 0,24 відповідно, тому була обрана як найточніша мережа серед 5 побудованих.

Рекомендація щодо зменшення невизначеності при проектуванні та вибору кількості прихованих шарів. Для встановлення прийнятого рівня узагальнюючої здатності мережі необхідно періодично тестувати мережу на незалежній контрольній множині даних і зупинити тренування, як тільки похибка цього тестування почне зростати. Однак, при цьому збільшується як час тренування, так і необхідна кількість даних. Кількість прихованих шарів нейронної мережі можна скоригувати і в процесі роботи саме тут потрібно провести ряд експериментів, щоб визначити найоптимальніший варіант.

Критерій оцінки якості функціонування мережі. Критерій оцінки слід поділити на внутрішні та зовнішні. Внутрішні критерії формуються на основі інформації множини даних, які були використані для навчання, тоді як зовнішні критерії використовують нову інформацію тестової множини, елементи яких не використовувались при навчанні. Оптимальна складність моделі мережі встановлюється за сукупністю зовнішніх та внутрішніх критеріїв.

До зовнішніх критеріїв вибору моделі можна віднести:

Критерій регулярності – зовнішній критерій Δ_i , що обумовлює невизначеність вибору моделі, обчислюється за тестовою множиною даних:

$$\Delta_i = \sum_{i=1}^{N_{np}} (\Phi_i^* - \Phi_i)^2 / \sum_{i=1}^{N_{np}} (\Phi_i^*)^2 \rightarrow \min, \quad (2)$$

де i – номер відгуку; Φ_i – значення відгуків, що продукуються мережею; Φ_i^* – точні значення відгуків.

Фізичний сенс застосування критерію регулярності полягає у виборі моделі, яка буде максимально точною на елементах тестової множини, які не входили до складу навчальної множини, тобто мали б найменшу невизначеність.

Так, наприклад, в табл. 2 представлено результат побудови п'яти кращих ШНМ за критерієм регулярності. Похибка на навчальній множині не перевищує 0,04, проте похибка при перевірці на тестовій множині, обчислена за формулою (2) значно більша.

Таблиця 1

Порівняльна характеристика похибок при побудові різних типів мереж

№	Тип мережі	Похибка обчислена на навчальній множині даних	Похибка обчислена на тестовій множині даних	Кількість вхідних нейронів	Кількість нейронів у прихованому шарі №1	Кількість нейронів у прихованому шарі №2
1	Радіально-базисна	8,34	8,87	17	5	0
2	Мережа загальної регресії	6,13	6,35	17	56	18
3	Багатошаровий перцептрон	0,19	0,24	17	0	0
4	Багатошаровий перцептрон	0,21	0,24	17	14	14
5	Багатошаровий перцептрон	0,20	0,25	17	14	0

Таблиця 2

Результати побудови мережі за критерієм регулярності
Мережа № 2

Фактичні дані (Φ^*)		Прогнозовані дані (Φ)		$(\Phi^* - \Phi)$		$(\Phi^* - \Phi)^2$	
к.т.	т	к.т.	т	к.т.	т	к.т.	т
10	1	10,42785	0,716658	-0,42785	0,283342	0,18306	0,080283
12	0	11,92325	-0,04612	0,076749	0,046121	0,00589	0,002127
12	0	11,85062	-0,01521	0,14938	0,015207	0,022314	0,000231
12	1	11,92977	0,0466	0,070235	0,9534	0,004933	0,908971
12	0	11,80585	0,037108	0,394152	-0,03711	0,155356	0,001377
7	1	11,02616	0,496316	-4,02616	0,503684	16,20998	0,253697
8	1	11,65793	0,139561	-3,65793	0,860439	13,38045	0,740356
7	1	11,56415	0,016027	-4,56415	0,983973	20,83142	0,968203
12	0	11,2358	0,201346	0,764198	-0,20135	0,583998	0,04054
9	1	8,283132	1,094666	0,716868	-0,09467	0,5139	0,008962
1063	6				\sum	51,8913	3,004747
					Δ	0,048816	0,500791

Рекомендації щодо зменшення невизначеності за критерієм регулярності полягає у виборі моделі, яка буде максимально точною на елементах тестової множини, які не входили до складу навчальної множини

Рекомендації щодо зменшення невизначеності за критерієм мінімального зміщення вимагає максимального співпадіння значень вихідної величини для двох моделей, де в якості навчальних елементів були використані дані різних підмножин навчальної множини. Критерій мінімального зміщення дозволяє обирати модель, яка "слабо реагує" на зміну навчальної множини і дозволяє вирішити задачу відновлення закону, що діє для зашумлених тестових даних.

Рекомендації щодо зменшення невизначеності за критерієм зміщеності показників в часі (допомагає оцінити рівень взаємозв'язку змінних). Окремі показники можуть мати різну післядію, тому розділене прогнозування кожного з них може забезпечити кращий результат. Відповідно, при наявності тісного взаємозв'язку між показниками, прогнозування їх сукупності покращує результати прогнозу кожного з них. Можливе також включення нового параметра, який може бути або додатковою ознакою або лінійною комбінацією вже включених ознак. Застосування цього критерію допомагає в оптимальному підборі таких ознак явища, які можуть забезпечити вищу точність прогнозування.

Рекомендації щодо зменшення невизначеності за критерієм фізичної достовірності – вимагає виключення моделей, які під час проведення експерименту можуть продукувати нереальні результати (великі викиди для множини, що прогнозується).

Навчання мережі. При навчанні мережі виникають три типи похибок. Перший тип – похибка одиничного вихідного елемента, яка враховується при реалізації процедури зворотного поширення похибки. Другий тип – похибка всієї мережі при конкретному вхідному сигналі, яка надає інформацію про те, наскільки правильним є відповідь мережі в даний момент часу. Третій тип – середня похибка мережі, обчислена після пред'явлення всього набору тренувальних даних, яка показує наскільки добре мережа засвоїла закономірності набору тренувальних даних. Остання похибка представляє собою усереднене по всьому набору тренувальних даних значення похибки другого типу.

Оскільки другий тип похибки часто обчислюється як відстань за модулем, то значення цієї похибки завжди позитивно і її усереднення не вимагає обчислення абсолютних значень.

На додаток до двох рівнів міри похибок, описаних вище, існують кілька індивідуальних мір похибок, які можуть бути використані. Кожна з них визначається конкретними вимогами до мережі. Найчастіше використовується міра похибки у вигляді простої різниці між значенням еталона та зна-

ченням відповідного виходу мережі для кожного вхідного значення з тренувальної послідовності:

$$b = t_i - a_i,$$

де t_i – еталонне значення для i -го виходу, a_i – поточне значення того ж виходу.

Для обчислення похибки всього вихідного шару можна трактувати еталони і виходи як точки багатовимірного простору і обчислювати відстань між ними. Відстань між двома точками обчислюється як квадратний корінь з суми квадратів різниць координат.

Похибку виходу мережі з N вихідними елементами можна обчислити за виразом (2). Така похибка відома як квадратична похибка (RSE - root squared error), що характеризує стандартну невизначеність:

$$u(b) = \sqrt{\sum_{i=1}^N (t_i - a_i)^2}. \quad (2)$$

Один з різновидів RSE - weighted RSE - зважена похибка (під терміном "ваги" в даному випадку розуміється не значення ваг мережі, а міра важливості, привласнена кожному з тренувальних прикладів):

$$b_i = W(t_i - a_i).$$

З кожним прикладом тренувального набору даних пов'язується вага W , яка вказує на значимість даного прикладу. Значення ваг повинні лежати між 0 і 1. RSE визначається в цьому випадку за формулою:

$$u(b) = W \sqrt{\sum_{i=1}^N (t_i - a_i)^2}.$$

Для завдань категоризації можна встановити поріг, так, щоб виходи вихідних елементів отримували значення або 0 (не належить даному класу), або 1 (належить даному класу). Для обчислення міри похибки недостатньо знати "так" чи "ні", необхідно обчислити розмір відхилення.

У деяких випадках представляється необхідним визначити похибку в термінах конкретного завдання, для вирішення якого сконструйовано мережу.

Як правило, після навчання нейромережі здійснюють контрольне відтворення даних, які склали навчальну множину. Якщо точність відтворення задовільна і відхилення знаходяться в допустимих межах, вважають, що побудовано задовільну модель і слід очікувати достатню якість відображення [2].

Причини збільшення невизначеності вихідних даних розробленої ШНМ можуть бути викликані:

– наявністю неточних даних з великою випадковою складовою. Для усунення цього явища підвищують вимоги до точності вимірювань; у випадку часового ряду можливе зменшення кроку дискретизації, наприклад, використання щомісячних значень замість річних;

– неврахуванням суттєвих ознак, які в значній мірі визначають закономірність; ця проблема може бути вирішена розширенням набору ознак, які приймаються до уваги.

Після отримання передбачених значень при наявності правильних можливо отримати абсолютні та відносні відхилення на всій контрольній множині, для кожного кроку прогнозування. При наявності задовільних результатів прогнозування на контрольній множині, можна вважати, що налаштована мережа для даної задачі має оптимальну складність і готова до відтворення даних, для яких немає відповідних відомих відгуків.

Рекомендації щодо зменшення невизначеності при навчанні нейромереж:

1. З навчальної вибірки береться поточний приклад (спочатку – перший) і його вхідні параметри (що представляють у сукупності вектор вхідних сигналів) подаються на вхідні синапси нейромережі, що навчається. Зазвичай кожен вхідний параметр прикладу подається на один відповідний вхідний синапс.

2. Нейромережа виконує задану кількість тактів функціонування, при цьому вектор вхідних сигналів розповсюджується по зв'язках між нейронами (пряме функціонування).

3. Вимірюються сигнали, видані тими нейронами, які вважаються вихідними.

4. Проводиться інтерпретація виданих сигналів, і обчислюється оцінка, що характеризує відмінність між виданою мережею відповіддю і необхідною відповіддю, наявною в прикладі. Чим менша оцінка, тим краще розпізнаний приклад, тим ближче видана мережею відповідь до потрібної. Оцінка, рівна нулю, означає що необхідна відповідність обчисленої і відомої відповідей досягнута. Відмітимо, що тільки що ініціалізована нейромережа (ненавчена), може видати правильну відповідь тільки абсолютно випадково.

5. Якщо оцінка прикладу дорівнює нулю, нічого не робиться. Інакше на підставі оцінки обчислюються поправочні коефіцієнти для кожної синаптичної ваги матриці зв'язків, після чого проводиться підстроювання синаптичних ваг (зворотне функціонування). У корекції ваг синапсів і полягає навчання.

6. Здійснюється перехід до наступного прикладу і перелічені вище операції повторюються. Прохід по всіх прикладах навчальної вибірки з першого по останній вважається одним циклом навчання.

При проходженні циклу кожен приклад має свою оцінку. Обчислюється, крім того, сумарна оцінка множини всіх прикладів навчальної вибірки. Якщо після проходження декількох циклів вона рівна нулю, навчання вважається закінченим, інакше цикли повторюються.

Невизначеність, що обумовлена похибками мережі по відношенню до розміру навчальної послідовності. Функціонування мережі поліпшується у міру ітеративного налаштування (тобто поступового наближення) вагових коефіцієнтів. Рівняння

$$\omega \leq p \leq \omega \log_2 \omega$$

описує результуючі межі числа ваг в мережі.

Вважаючи границю похибок $0 \leq \epsilon \leq 1/8$ [3], кількість тренувальних прикладів має бути приблизно рівною кількості ваг мережі, помноженому на зворотну величину похибки. Наприклад, для граничної похибки $\epsilon = 0,1$ необхідно використовувати навчальну послідовність n в 10 разів більшу кількості ваг ω . Ця залежність описується формулою:

$$n \geq \frac{\omega}{\epsilon}$$

Причина того, що значення похибки відіграє значну роль, пов'язана із співвідношенням між узагальнюючою здатністю і точністю.

Рекомендації по зменшенню невизначеності за кількістю прихованих елементів. При виборі кількості прихованих елементів h , по-перше, не варто h більше, ніж подвійну кількість вхідних елементів, тобто $h \leq 2i + 1$, по-друге, якщо необхідно забезпечити високу узагальнюючу здатність, необхідно обирати значно меншу кількість прихованих елементів. Крім того, кількість навчальних даних має бути що найменше в $1/\epsilon$ разів більше кількості ваг у мережі.

Якщо розмірність даних може бути зменшена, то слід використовувати меншу кількість прихованих одиниць.

При навчанні на безструктурних входах необхідно, щоб кількість прихованих одиниць була більшою, ніж кількість входів.

Кількість прихованих одиниць, необхідних для вирішення задач класифікації, має збільшуватися зі збільшенням кількості класів, точніше, кількості областей, на які має бути розділено простір входів, що є більш важливим фактором, ніж кількість класів само по собі.

Функції активації елемента (нейрона) підсумовує зважені входи від усіх приєднаних елементів і стискає їх в заданий діапазон значень. Цей діапазон вибирається рівним або $(0; -1)$, або $(-1; +1)$. При цьому великі значення завжди стискаються таким чином, щоб вони вносили зменшуваний внесок. З цієї причини, а також тому, що мережа повинна моделювати нелінійні процеси, функції активації повинні бути нелінійними.

Рекомендації щодо зменшення невизначеності: для одержання якісних результатів навчальна, контрольна і тестова множини повинні бути репрезентативними з погляду суті задачі. Більш того, ці множини повинні бути репрезентативними кожна окремо. Якщо вибірка навчальних даних $n_{\text{нав}}$ не репрезентативна, то модель, як мінімум, буде не дуже точною, а в гіршому випадку – непридатною.

Показник точності прогнозування E оцінюється сумою модулів похибок, що обумовлюють невизначеність при проектуванні та навчанні нейронної мережі:

$$E = \frac{\sum |e_i|}{n_{\text{виб}}}$$

Таблиця 3

Порівняльна характеристика нейромереж

№	Тип мережі	Похибка навчання	Вибіркова похибка	Похибка тестування	Кількість вхідних нейронів	Кількість нейронів у прихованому шарі №1
1	RBF 17-17-3-2-2	0,218990	0,273355	0,782891	17	3
2	Linear 1:1-2-2	0,399745	0,166594	0,276304	1	0
3	MLP 1:1-5-2-2	0,114087	0,035489	0,300058	1	5
4	MLP 2-2-4-2-2	0,013125	0,015855	0,542260	2	4
5	GRNN 17-17-6-3-2-2	0,002467	0,000000	0,043455	17	6

Чим нижче значення E , тим менша невизначеність результату прогнозу.

Таким чином, за результатами даної серії експериментів оптимальною за показниками точності визначена мережа, що показала найліпший результат серед 186 побудованих нейромереж. Враховуючи сформульовані рекомендації до зменшення невизначеності при її побудові використано правило навчання з вчителем, сигмоїдальна функція активації, два прихованих шари та 9 нейронів у них. Порівняння результатів прогнозування нейромережами, наведено у табл. 3.

Найменшу похибку при перевірці на тестовій множині показала мережа № 5, яка має при навчанні на тестовій множині похибку 0,043, а при перевірці на тестовій множині похибку 0,001.

Висновки

В статті проаналізовано підходи до побудови ШНМ, що забезпечить найменшу невизначеність прогнозу діагностичних ознак серцево-судинної системи. Виявлено джерела невизначеності, що виникають як на етапі введення вхідних даних, так і на етапах побудови та навчання мережі.

Запропоновано рекомендації щодо зменшення невизначеності прогнозу для кожного з етапів побудови моделі. В якості прикладу наведено побудовану найефективнішу за показником точності архітектуру штучної нейронної мережі, що містить багатшаровий перцептрон, імовірнісну нейронну мережу, мережу зворотного поширення похибки та їх модифікації.

Побудована нейромережа може бути використана в кардіологічному відділенні для прогнозування виникнення мікроінфаркту міокарда як така, що надає прогноз з найменшою невизначеністю, бо має найменшу помилку тестової множини, а отже високу здатність до навчання та узагальнення.

Список літератури

1. Кошечкина Л.А. Особенности оценивания неопределенности химико-аналитических измерений в медицине / Л.А. Кошечкина // Системи обробки інформації: зб. наук. пр. – Х.: ХУПС, 2008. – Вип. 4 (71). – С. 102-104.
2. Kevin Swingler. Applying Neural Networks. A practical Guide) [Електронний ресурс] / Kevin Swingler. – Режим доступу до ресурсу: <http://www.matlab.exponenta.ru/neuralnetwork>.
3. Baum E.B. What net size gives valid generalisation / E.B. Baum, D. Haussler // Neural Computation – 1989. – 1 (1). – P. 151-160.

Надійшла до редколегії 17.01.2012

Рецензент: д-р техн. наук, проф. Ю.П. Мачехин, Харківський національний університет радіоелектроніки, Харків, Україна.

ПУТИ УМЕНЬШЕНИЯ НЕОПРЕДЕЛЁННОСТИ ПРОГНОЗА СОСТОЯНИЯ ОРГАНИЗМА ЧЕЛОВЕКА ПРИ НЕЙРОСЕТЕВОМ МОДЕЛИРОВАНИИ

М.Ю. Буриченко, О.Б. Иванец

В работе проведен анализ источников неопределенности, возникающих при построении моделей для решения задач прогнозирования диагностических признаков биологических объектов с помощью искусственных нейронных сетей. Предложены рекомендации по уменьшению неопределенности прогноза на разных этапах построения нейросети.

Ключевые слова: неопределенность, искусственные нейронные сети, модель прогнозирования, медико-биологические параметры, диагностические признаки.

WAYS OF REDUCING FORECAST UNCERTAINTY IN THE STATE OF THE HUMAN NEURAL NETWORK MODELING

M.Yu. Burichenko, O.B. Ivanets

The paper analyzes the sources of uncertainty arising from the construction of models for solving problems of prediction of the diagnostic features of biological objects by means of artificial neural networks. Recommendations for reducing forecast uncertainty at various stages of building a neural network.

Keywords: uncertainty, artificial neural networks, model forecasting, medical and biological parameters, diagnostic signs.