

УДК 629.735.083.03.004.58:004.801.3(045)

О.С. Якушенко, к.т.н., с.н.с.

ОПТИМІЗАЦІЯ РОЗМІРУ НЕЙРОННОЇ МЕРЕЖІ ДІАГНОСТУВАННЯ ТУРБОРЕАКТИВНОГО ДВОКОНТУРНОГО ДВИГУНА З УРАХУВАННЯМ ЕФЕКТУ ПЕРЕНАВЧАННЯ

Національний авіаційний університет
E-mail: yuysss@mail.ru

Розглянуто метод оптимізації розміру нейронної мережі, призначеної для класифікації технічного стану турбореактивного двоконтурного двигуна. Для визначення моменту виникнення ефекту перенавчання нейронної мережі використано метод, який ґрунтується на порівнянні зміни параметра якості розпізнавання, отриманому під час розпізнавання навчального та контрольного наборів даних. Показано, що під час виникнення ефекту перенавчання мережі якість розпізнавання навчального набору у процесі навчання продовжує покращуватися, а при розпізнаванні контрольного набору даних починає падати. Проведено оптимізацію дво-, три- та чотиришарових мереж. Наведено отримані результати.

газотурбінний двигун, діагностування, кількість нейронів, нейронна мережа, оптимізація структури

Постановка проблеми

У роботах [1–5] розглянуто методи первинної оптимізації нейронної мережі, пов'язані з отриманням даних для навчання таких мереж, вибором методу навчання, структури мережі та типу її нейронів. Метод скінченної оптимізації, спрямований на визначення розміру мережі, забезпечить високу якість класифікації технічного стану (ТС) за необхідного рівня узагальнення. Розроблюваний метод має враховувати можливості виникнення ефекту перенавчання. Із виникненням цього ефекту мережа втрачає властивість до узагальнення і починає строго описувати дані навчального набору.

Аналіз досліджень і публікацій

Перспективним напрямком автоматизованого визначення класу ТС газотурбінного двигуна (ГТД) є використання нейронних мереж. Однією з особливостей такої мережі є гнучкість і поліморфізм її структури.

У роботах [1–5] розглянуто питання вибору методу навчання нейронної мережі та визначення її структури і типу нейронів. У роботі [4] викладено дані про визначення кількості нейронів у шарах нейронної мережі, але не наведено закінченого алгоритму. У літературі рекомендацій щодо оптимізації немає або вони мають загальний характер [6].

Мета роботи – створення методу визначення кількості нейронів у проміжних шарах нейронної мережі з урахуванням можливості виникнення ефекту перенавчання. Як об'єкт досліджень обрано статичні нейронні мережі.

Явище перенавчання нейронної мережі

Більшість методів класифікації, такі, як дискримінантний аналіз, метод мінімізації ризику та інші для виконання обчислень використовують результати статистичної обробки (математичні сподівання, дисперсії), а не навчальну вибірку. Це дозволяє значною мірою нівелювати вплив випадкової складової, яку можна розділити на три частини:

– розсіювання даних про центр класу, пов'язане з реальною різницею між зовнішніми проявами ТС об'єктів, які належать до одного класу;

– розсіювання, пов'язане з похибками вимірювання параметрів, які використано під час класифікації;

– неповна адекватність методів розрахунку значень ознак ТС об'єкту.

У зв'язку з цим треба відзначити дві особливості нейронних мереж.

Першою особливістю нейронних мереж є те, що вони на відміну від більшості інших методів у процесі навчання (настроювання параметрів) спираються на індивідуальні розрахункові точки, а не на результати статистичної обробки навчальних даних.

Другою особливістю такої мережі є те, що її архітектура поліморфна, може бути легко змінена. При цьому верхня межа потужності мережі (кількість шарів та кількість нейронів у кожному шарі) лімітується фактично тільки характеристиками наявної обчислювальної техніки та часом навчання.

Наведені особливості нейронних мереж призводять до виникнення так званого ефекту перенавчання нейронної мережі. За його виникнення мережа починає описувати неусереднені дані для класів, а саме навчальну вибірку. Інші дані, взяті з тієї ж генеральної сукупності класифікуються неправильно.

Для виявлення ефекту перенавчання використовують механізм контрольної перевірки [6]. Застосовуючи цей метод, частину навчальних спостережень резервують як контрольні і не використовують для навчання мережі. Під час навчання ці спостереження застосовують для незалежного контролю якості навчання. Спочатку помилки мережі на навчальній і контрольній множині будують майже однаковими. Якщо вони суттєво відрізняються, то, ймовірно, це означає, що розбиття спостережень на дві множини не забезпечило їх однорідності. У процесі навчання мережі помилка зменшується, і, поки навчання зменшує функцію помилок, помилка на контрольній множині також буде зменшуватися. Якщо ж контрольна помилка перестала зменшуватися або стала зростати, це вказує на те, що мережа почала набагато більше відповідати навчальним даним і навчання варто припинити.

У цьому разі варто зменшити кількість нейронів або шарів, оскільки мережа надто потужна для вирішення цього завдання.

Якщо навпаки мережа має недостатню потужність, щоб відтворити наявну залежність, то явище перенавчання спостерігатися не буде й обидві помилки – навчання і перевірки – не досягнуть досить малого рівня.

Проблеми відшукування глобального мінімуму або вибору розміру мережі, що виникають у процесі робіт з нейронними мережами, призводять до того, що під час практичної роботи доводиться експериментувати з великою кількістю мереж різних конфігурацій, часом навчаючи кожну з них декілька разів і порівнюючи отримані результати.

Головним критерієм вибору в цих випадках є контрольна похибка. При цьому застосовується правило, відповідно до якого з двох нейронних мереж із приблизно однаковими контрольними похибками варто вибирати простішу.

Необхідність багатократних експериментів веде до того, що контрольна множина починає відігравати ключову роль у виборі моделі нейронної мережі, тобто стає частиною процесу навчання. Тим самим її роль як незалежного критерію якості навчання послаблюється, оскільки за великої кількості експериментів виникає ризик перенавчання нейронної мережі на контрольній множині. Для того щоб гарантувати надійність мережі резервують ще тестову множину спостережень. Скінченна мережа тестується на даних із цієї множини, щоб переконатися, що результати, які досягнуті на навчальній і контрольній множині, є реальними. Для того щоб добре виконувати свою роль, тестова множина має використовуватися тільки один раз. Якщо її використовувати повторно для коригування процесу навчання, то вона фактично перетвориться в контрольну множину.

Отже, процедура побудови нейронної мережі складається з таких кроків:

- обрання початкової конфігурації мережі;
- навчання мережі і моделювання з оцінкою контрольної помилки;
- виявлення ефекту перенавчання і коригування конфігурації мережі.

Для формування цих множин необхідно пам'ятати, що кожна з них є класифікованим набором даних, який містить параметри об'єкту, що належать до всіх шуканих класів ТС.

Наведені вибірки мають, якщо йдеться про ГТД, складатися з параметрів, що можуть бути використані для визначення ТС ГТД, та з інформації, що формально характеризує (описує, визначає) клас ТС, до якого віднесено двигун.

Кожна така вибірка складається з ряду записів (точок), а кожен запис – з параметрів, що характеризують ТС та, можливо, режими та умови роботи об'єкта. Під час цього у складі вибірки необхідно мати хоча б по декілька записів, що характеризують кожен клас ТС. У разі формування динамічної мережі, яка визначає клас ТС за результатами аналізу зміни параметрів функціонування ГТД у часі, необхідно мати вже не набір записів, а набір груп записів, кожна з яких характеризує певний клас (або групу класів) ТС.

Перевірка наявності ефекту перенавчання для даних розпізнавання класу технічного стану

Контрольна вибірка, яка необхідна для перевірки мережі на перенавчання, готується аналогічно навчальній за методом, наведеним у роботі [1] і відрізняється тільки тим, що випадкову складову при її формуванні взято з іншої частини тієї ж генеральної сукупності.

Як і в роботах [1; 2; 4], за параметр адекватності мережі вхідним даним використано відсоток грубих помилок розпізнавання Δ^r . Докладнішу характеристику та метод обчислення цього параметра наведено в роботі [1]. Параметр, отриманий за контрольною вибіркою, будемо позначати як Δ_k^r .

Як об'єкт контролю обрано двигун ПС-90А [7]. Набір діагностичних даних для навчання та контролю мережі містив інформацію про 16 класів ТС. Опис класів та механізм отримання даних наведено в роботі [1].

Для підтвердження наявності ефекту перенавчання, аналогічного описаному, для випадку класифікації ТС ГТД провели навчання заздалегідь перерозмірених нейронних мереж.

Проведено навчання чотиришарової мережі з кількістю нейронів у шарах 7–20–20–16 та 7–40–50–16. Як функцію активації вибрано гіперболічний тангенс. Навчання велося за алгоритмом CGP до 3500 циклів. Результати, отримані після 3500 циклів, наведено у табл. 1.

Таблиця 1

Результати навчання мереж різного розміру

Мережа	Помилка розпізнавання набору Δ^r , %	
	навчального	контрольного
7–20–20–16	0,3	16
7–40–50–16	0	21,9

Порівняльний аналіз результатів розпізнавання навчального та контрольного наборів даних, наведених у табл. 1, показує, що обидві мережі навчилися добре (майже на 100 % правильно) розпізнавати навчальний набір даних, але ні одна не змогла досить точно розпізнати контрольний набір даних. Це свідчить про те, що мережа перерозмірена та перенавчена та (або) про те, що навчальний набір не досить повно охоплює всі можливі випадки ТС (вибірki недостатньо представницькі).

Після виконання кожних 500 циклів навчання мережі переривалося і проводилася оцінка якості розпізнавання навчального та контрольного наборів даних. Отримані результати показано на рис. 1.

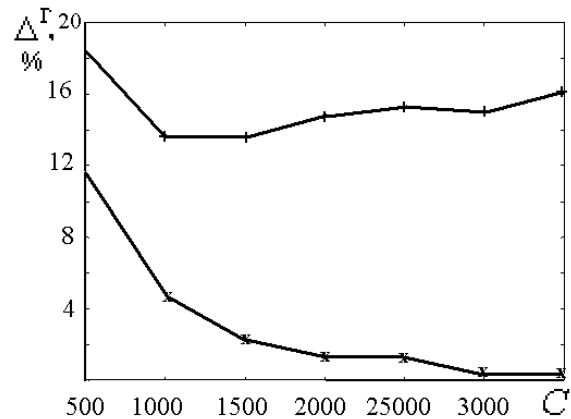


Рис. 1. Залежність величини Δ^r для навчального (x) та контрольного (+) наборів від кількості циклів навчання (мережа 7–20–20–16)

Як можна бачити з цих даних на початковому етапі навчання мережі (перші 1000 циклів) обидва параметри змінюються синфазно. Далі величина параметра Δ^r для навчального набору постійно зменшується і наприкінці майже досягає нульового значення. Для параметра Δ^r контрольного набору між 1000 та 1500 циклами настає насичення і в разі подальшого навчання цей параметр починає збільшуватися. Цю обставину можна використати для знаходження моменту припинення навчання. Тому необхідно відстежити момент початку сталого зростання цього параметра.

Проведення оптимізації потужності нейронної мережі

Характер залежностей на рис. 1 може бути результатом того, що навчальний набір не досить повно охоплює усі можливі випадки ТС (вибірki недостатньо представницькі). У цих вибірках кожен клас ТС було репрезентовано 20 точками. Тому було прийнято рішення збільшити подання кожного класу в навчальному та контрольному наборах до 200 точок (усього 3600 точок у наборі).

Загальну характеристику отриманого навчального набору наведено у табл. 2. На рис. 2 показано зміну отриманих діагностичних відхилень у класах.

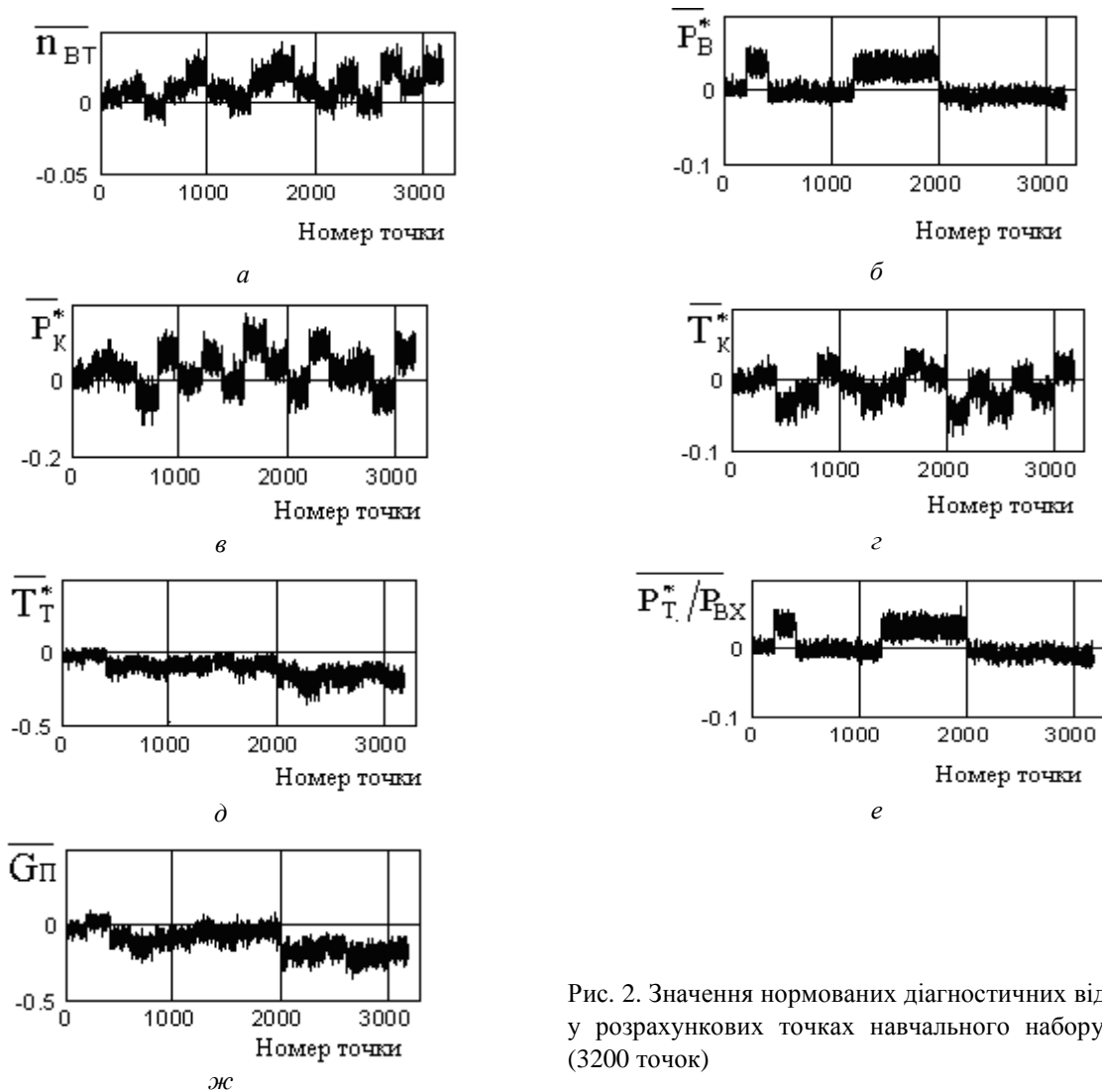


Рис. 2. Значення нормованих діагностичних відхилень у розрахункових точках навчального набору даних (3200 точок)

Таблиця 2

Загальна характеристика розширеного навчального набору даних

Параметр	Базове значення (середньостатистичний ГТД)	Діапазон зміни
Частота обертання, %: вентилятора КВТ	90 93,99	– 89,9–95,5
Повний тиск, МПа: за вентилятором за КВТ	1,69 2790	1,59–1,73 22,9–31,1
Повна температура, К: за КВТ за турбіною	808 820	771–871 795–1113
Відношення повного тиску за турбіною до тиску на вході до ГТД	1,568	1,47–1,61
Витрата палива, кг/г	5564	5608–7533

У ході оптимізації розміру мережі було проведено навчання дво-, три- та чотиришарової мережі з різною кількістю нейронів у шарі. У двошаровій мережі змінювалась кількість нейронів у вхідному шарі. У тришаровій мережі кількість нейронів у вхідному шарі була 7, у чотиришаровій мережі 7 та 14, а кількості нейронів у проміжних шарах (параметр, що оптимізується) були однаковими. У всіх випадках кількість вихідних нейронів дорівнювала 16.

Як функцію активації на цьому етапі вибрано гіперболічний тангенс.

За результатами цих експериментів побудовали залежності мінімального значення відсотка грубих помилок розпізнавання контрольного набору від кількості нейронів та кількості циклів навчання, необхідних для цього (рис. 3–6).

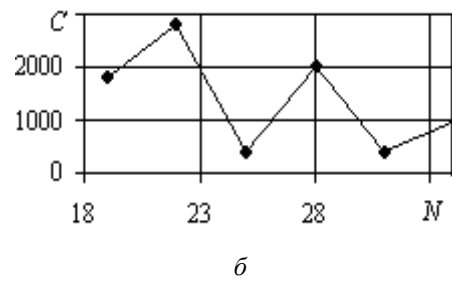
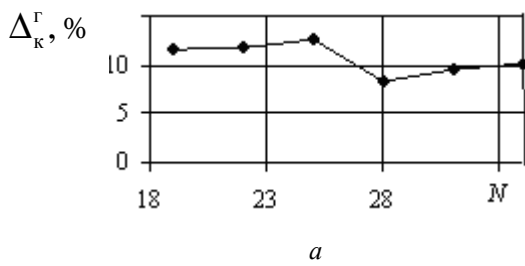


Рис. 3. Залежність параметра Δ_k^Γ (а) та кількості циклів навчання C (б) від кількості нейронів N у вхідному шарі двошарової мережі

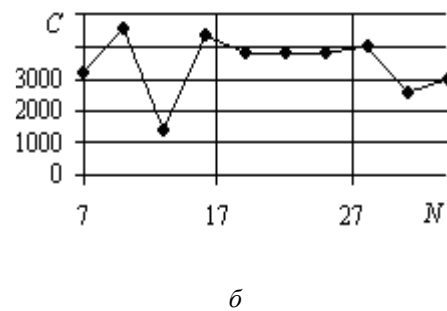
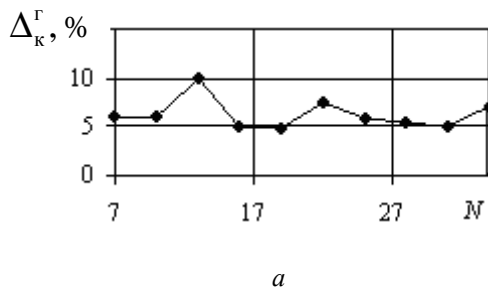


Рис. 4. Залежність параметра Δ_k^Γ (а) та кількості циклів навчання C (б) від кількості нейронів N у проміжному шарі тришарової мережі

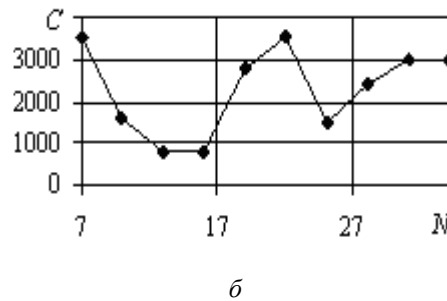
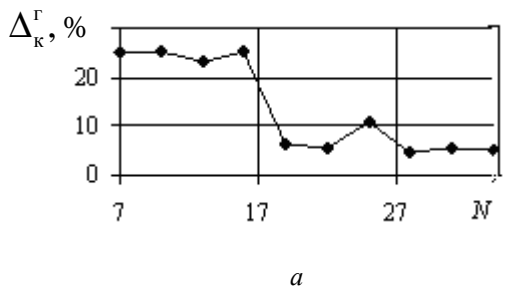


Рис. 5. Залежність параметра Δ_k^Γ (а) та кількості циклів навчання C (б) від кількості нейронів N у проміжних шарах чотиришарової мережі при 7 нейронах у вхідному шарі

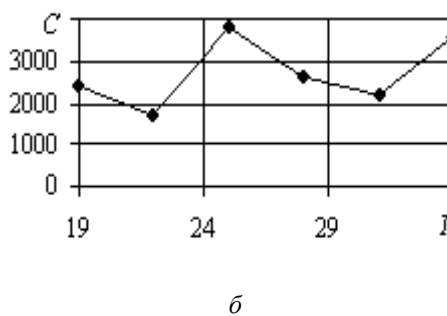
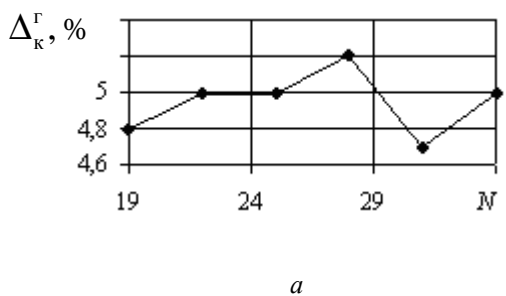


Рис. 6. Залежність параметра Δ_k^Γ (а) та кількості циклів навчання C (б) від кількості нейронів N у проміжних шарах чотиришарової мережі при 14 нейронах у вхідному шарі

Двошарові мережі (рис. 3), як і у випадку мереж Кохонена [1], не дозволяють отримати необхідний рівень точності розпізнавання, забезпечуючи значення параметра Δ_k^r понад 8 %.

Тришарові мережі забезпечують значення параметра Δ_k^r близько 5 %, але необхідно декілька спроб навчання мережі для отримання такого рівня.

Чотиришарова мережа забезпечує трохи нижчий рівень параметра Δ_k^r і більш стабільні результати навчання.

Додаткову стабілізацію процесу навчання можна досягнути, збільшивши кількість нейронів у вхідному шарі.

Величина Δ_k^r для тришарової мережі при $N > 16$ майже не залежить від кількості нейронів, стабілізуючись на рівні 5 %.

Величина Δ_k^r для тришарової мережі у діапазоні N від 7–16 трохи вища за 5 %, але може бути знижена проведенням ряду спроб навчання мережі та вибором мережі, що забезпечує кращий результат.

Оптимальна кількість нейронів проміжних шарів у чотиришаровій мережі лежить приблизно у межах 16–19.

Кількість циклів навчання не залежить практично від кількості нейронів.

До недоліків чотиришарової мережі слід віднести тривалий термін навчання – 8–12 год на комп'ютері з тактовою частотою 900 МГц та ємність оперативної пам'яті 400 Мбайт.

Усі наведені результати отримано для мережі з функцією активації нейронів типу гіперболічного тангенса. Можливо також використовувати логістичну функцію, тому було проведено оптимізацію тришарової мережі, нейрони якої використовують цю функцію. Узагальнені отримані результати показано на рис. 7.

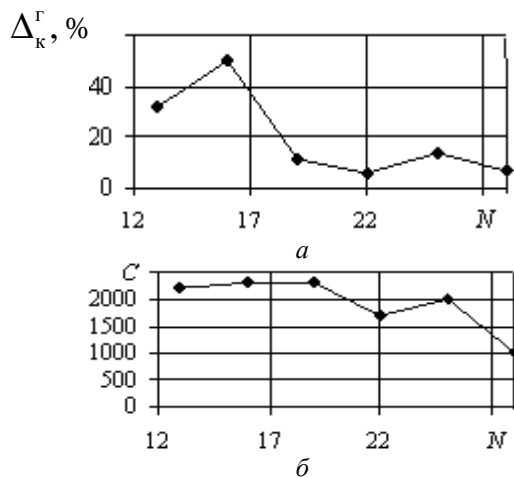


Рис. 7. Залежність параметра Δ_k^r (а) та кількості циклів навчання C (б) від кількості нейронів N у проміжному шарі тришарової мережі з логістичною функцією активації

З наведених даних можна бачити, що мережа з такими нейронами забезпечує приблизно ту саму точність розпізнавання (більше 6 %), що й з нейронами, які використовують функцію гіперболічного тангенса за оптимальної кількості нейронів 19–22. Якщо кількість менша за 19 якість класифікації різко знижується. Отже, більш доцільно використовувати як функції активації функції гіперболічного тангенса.

Оскільки значення Δ_k^r у три- та чотиришаровій мережах майже однакові, більш доцільно використовувати мережу з меншою кількістю нейронів, тобто тришарову мережу (рис. 8).

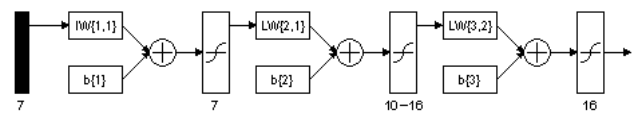


Рис. 8. Оптимальна структура мережі

Висновки

Розроблено метод визначення оптимального розміру нейронної мережі часу закінчення процесу її навчання. Метод ґрунтується на визначенні часу виникнення ефекту перенавчання мережі. Наведені дані показують досить високу ефективність методу розпізнавання ТС ГТД із використанням нейронних мереж, оптимізованих та навчених за таким методом (частка грубих помилок 4,5–5 %). Але результати було отримано з використанням даних, які не мають похибок та грубих промахів вимірювання, тому наступним етапом розвитку цього методу буде випробування його на стійкість за наявності у вхідних даних (навчальному наборі) таких похибок. Крім того, може виявитися доцільним визначати клас ТС з урахуванням режиму та умов роботи турбоактивного двоконтурного двигуна.

Література

1. Кучер О.Г. Розпізнавання технічного стану авіаційних ГТД з використанням нейронних мереж / О.Г. Кучер, О.С. Якушенко, В.Ю. Сухоруків // Вісник двигунобудування. – Запоріжжя: ЗНТУ, ХАІ, МоторСіч, 2002. – №1. – С. 101–106.
2. Якушенко О.С. Нейронні мережі для діагностування газотурбінних двигунів/ О.С. Якушенко // Вісник НАУ. – 2004. – № 2 (20). – С. 67–71.

3. *Многокритериальное* диагностирование проточной части авиационных газотурбинных двигателей/ А.Г. Кучер, С.А. Дмитриев, А.В. Попов, А.С. Якушенко // Вісник двигунобудування. – Запоріжжя: ОАО Мотор-Січ, 2009. – № 3. – С. 166–171.

4. *Первинна* оптимізація архітектури нейронної мережі/ А.Г. Кучер, О.С. Якушенко, В.М. Степаненко, В.Є. Мільцов // Вісник двигунобудування. – Запоріжжя: ОАО Мотор-Січ. – 2009. – № 3. – С. 166–171.

5. *Нейросетевое* распознавание технического состояния авиационных газотурбинных двигателей/ А.М. Пашаев, П.Ш. Абдуллаев, А.С. Якушенко, А.Д. Мирзоев // Изв. Нац. акад. наук Азербайджана. Информатика и проблемы управления. – 2008. – № 3 (28). – С. 3–7.

6. *Медведев В.С.* Нейронные сети. Matlab 6 / В.С. Медведев, В.Г. Потемкин. / под общ. ред. В.Г. Потемкина. – М.: ДИАЛОГ-МИФИ, 2002. – 496 с.

7. *Двигатель* ПС-90А. Руководство по технической эксплуатации. в 3 кн. – Пермь: ПНПО "Авиадвигатель", 1990.

Стаття надійшла до редакції 04.12.09.

А.С. Якушенко

ОПТИМИЗАЦИЯ РАЗМЕРА НЕЙРОННОЙ СЕТИ ДИАГНОСТИКИ ТУРБОРЕАКТИВНОГО ДВУХКОНТУРНОГО ДВИГАТЕЛЯ С УЧЕТОМ ЭФФЕКТА ПЕРЕУЧИВАНИЯ

Национальный авиационный университет

газотурбинный двигатель, диагностика, количество нейронов, нейронная сеть, оптимизация структуры

Одной из наиболее важных проблем при создании нейронных сетей является определение оптимального размера их структуры (количество слоев в сети и количество нейронов в каждом слое). Оптимизацию размера нейронных сетей, предназначенных для классификации технического состояния газотурбинных двигателей, проводили с учетом возможности возникновения эффекта переучивания, когда обучаемая сеть утрачивает свойство обобщения и начинается строго описывать учебный набор данных. Для определения момента, когда в обучаемой нейронной сети появляется эффект переучивания, использован метод трех выборок, основанный на сравнении изменений параметров, характеризующих качество распознавания данных. Для определения момента, когда качество распознавания контрольного набора данных начинает устойчиво снижаться, а качество распознавания учебного набора продолжает улучшаться, процесс обучения периодически прерывается и производится классификация учебного и контрольного наборов данных. Приведены результаты оптимизации двух-, трех- и четырехслойной нейронных сетей, сгенерирован расширенный учебный набор данных, описывающий 16 классов технического состояния двигателя, а каждый класс представлен 200 точками. В статье рассмотрен алгоритм оптимизации нейронной сети и результаты его использования. Полученная структура нейронной сети обеспечивает наилучшее качество распознавания класса технического состояния газотурбинного двигателя и высокий уровень обобщения сети при наименьшем размере самой сети.

Alexander S. Yakushenko

TURBOFAN ENGINE DIAGNOSTICS NEURON NETWORK SIZE OPTIMIZATION METHOD WHICH TAKES INTO ACCOUNT OVERLEARNING EFFECT

National Aviation University

amount of neurons, diagnostics, gas turbine engine, neuron network, optimization of structure

The article is devoted to the problem of gas turbine engine (GTE) technical state class automatic recognition with operation parameters by neuron networks. The one of main problems for creation the neuron networks is determination of their optimal structures size (amount of layers in network and count of neurons in each layer). The method of neuron network size optimization intended for classification of GTE technical state is considered in the article. Optimization is cared out with taking into account of overlearning effect possibility when a learning network loses property of generalization and begins strictly describing educational data set. To determinate a moment when overlearning effect is appeared in learning neuron network the method of three data sets is used. The method is based on the comparison of recognition quality parameters changes which were calculated during recognition of educational and control data sets. As the moment when network overlearning effect is appeared the moment when control data set recognition quality begins deteriorating but educational data set recognition quality continues still improving is used. To determinate this moment learning process periodically is terminated and simulation of network with education and control data sets is fulfilled. The optimization of two-, three- and four-layer networks is conducted and some results of optimization are shown. Also the extended educational set is created and shown. The set describes 16 GTE technical state classes and each class is represented with 200 points (200 possible technical state class realizations) instead of 20 points using in the former articles. It was done to increase representativeness of data set. In the article the algorithm of optimization is considered and some results which were obtained with it are shown. The results of experiments were analyzed to determinate most optimal neuron network structure. This structure provides most high-quality GTE technical state classification and high level of network generalization.