

МЕТОДЫ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РАСЧЕТОВ НА ЭВМ ТЕПЛОВЫХ, ГИДРОДИНАМИЧЕСКИХ И ФИЗИКО-ХИМИЧЕСКИХ ПРОЦЕССОВ

Часть 1. Методы нейронных сетей для расчета термодинамических свойств веществ.

Кандидат технических наук, Катковский Е.А.
E-mail: katkovsky@starnet.ru

Аннотация.

В части 1. работы приводится теоретическое обоснование и практические результаты применения нейронных сетей для определения и вычисления термодинамических свойств веществ.

На примере воды, находящейся в разных областях состояния, приведены конкретные результаты применения нейронных сетей ряда классов для группового расчета её термодинамических свойств. Исследованы минимально необходимые размеры нейронных сетей в зависимости от числа определяемых термодинамических свойств. Проанализирована точность получаемых результатов.

Введение.

На применение нейронных сетей авторов натолкнула теорема академика Колмогорова, завершившая его серию исследований для непрерывных функций: «*Каждая непрерывная функция N переменных, заданная на единичном кубе N -мерного пространства, представима с помощью операций сложения, умножения и суперпозиции из непрерывных функций одного переменного*» [1]. На этом фоне совершенно неожиданно выглядит тот факт, что любой многочлен от многих переменных может быть получен из одного произвольного нелинейного многочлена от одного переменного с помощью линейных операций и суперпозиции. Доказательство этой теоремы дано в [2].

Теория нейронных сетей включают широкий круг вопросов из разных областей науки: биофизики, математики, информатики, схемотехники и технологии. Поэтому понятие "нейронные сети" детально определить сложно.

Искусственные нейронные сети (ИНС) — совокупность моделей биологических нейронных сетей. **ИНС** представляют собой сеть элементов — искусственных нейронов (**перцептронов**) — связанных между собой синаптическими соединениями. **ИНС** обрабатывает входную информацию и в процессе изменения своего состояния во времени формирует совокупность выходных сигналов. Работа сети состоит в преобразовании входных сигналов во времени, в результате чего меняется внутреннее состояние сети и формируются выходные воздействия. Большинство моделей **ИНС** требуют обучения. В общем случае, **обучение** — такой выбор параметров сети, при котором **ИНС** лучше всего справляется с поставленной проблемой. Обучение — это задача многомерной оптимизации, и для ее решения существует множество алгоритмов.

Искусственные нейронные сети — набор математических и алгоритмических методов для решения широкого круга задач. Выделим характерные черты

LES METHODES DES RESEAUX NEURONIQUES POUR LES CALCULS ELECTRONIQUES DES PROCES THERMIQUES, HYDRODYNAMIQUES ET PHYSICO-CHIMIQUES

La partie 1. Les méthodes des réseaux neuroniques pour le calcul des propriétés thermodynamiques des substances.

Docteur ès sciences, Katkovsky E.A.
E-mail: katkovsky@starnet.ru

L'annotation.

Dans la partie 1. du présent article nous donnons l'argumentation théorique et les résultats pratiques de l'application des réseaux neuroniques afin de définir et de calculer des propriétés thermodynamiques des substances.

À l'exemple de l'eau se trouvant dans de différents domaines de l'état, nous citons les résultats concrets de l'application des réseaux neuroniques d'une série de classes pour faire le calcul de groupe de ses propriétés thermodynamiques. On a étudié les paramètres nécessaires au minimum des réseaux neuroniques en fonction de la quantité des propriétés thermodynamiques que l'on définit. On a fait l'analyse de l'exactitude des résultats reçus.

L'introduction.

L'application des réseaux neuroniques des auteurs était suggérée par le théorème de l'académicien Kolmogorov, celui qui a finalisé une série de ses études concernant des fonctions continues : «*Chaque fonction continue N des variables, donnée sur le cube unitaire de l'espace en N dimensions, peut être présentée avec l'aide des opérations de l'addition, de la multiplication et de la juxtaposition des fonctions continues d'une variable*» [1]. Dans cette situation le fait que tout polynôme de plusieurs variables peut être reçu d'un polynôme non linéaire choisi à volonté d'une variable avec l'aide des opérations linéaires et avec celle de la juxtaposition paraît contre toute attente. La preuve de ce théorème est donnée dans [2].

La théorie des réseaux neuroniques comprend un grand nombre de problèmes de différents domaines de la science : de la biophysique, des mathématiques, de l'informatique, des circuits intégrés et des technologies. C'est pourquoi il est difficile de définir la notion "les réseaux neuroniques" en détail.

Les réseaux neuroniques artificiels (RNA) – c'est l'ensemble des modèles des réseaux biologiques neuroniques. **RNA** représentent le réseau des éléments - les neurones artificiels («**perceptrones**») – qui sont liés entre eux par des liaisons synaptiques. **RNA** traite l'information d'entrée, et au cours du changement de son état dans le temps forme l'ensemble des signaux de sortie. Le travail du réseau comprend la transformation des signaux d'entrée dans le temps, à la suite de quoi l'état intérieur du réseau change et les actions de sortie se forment. La plupart des modèles **RNA** demandent l'enseignement. Au total, **l'enseignement** - c'est un choix des paramètres du réseau, avec lequel **RNA** s'acquitte le mieux du problème posé. L'enseignement – c'est l'enjeu de l'optimisation multidimensionnelle, et pour sa réalisation il y a une multitude d'algorithmes.

Les réseaux neuroniques artificiels – c'est une composition des méthodes mathématiques et algorithmiques créée afin de résoudre un grand nombre de problèmes. Nous sélectionnerons ci-dessous les traits distinctifs des **RNA** en tant que l'outil universel pour la résolution des problèmes :

искусственных **ИНС** как универсального инструмента для решения задач:

ИНС — средство обработки информации:

- а) гибкая модель для нелинейной аппроксимации многомерных функций;
- б) средство прогнозирования во времени для процессов, зависящих от многих переменных;
- в) классификатор по многим признакам, дающий разбиение входного пространства на области;
- г) средство распознавания образов;
- д) инструмент для поиска по ассоциациям;
- г) модель для поиска закономерностей в массивах данных.

ИНС свободны от ограничений обычных компьютеров благодаря параллельной обработке и сильной связанности нейронов.

Современные искусственные **ИНС** по сложности и "интеллекту" приближаются к нервной системе насекомого, но уже сейчас демонстрируют ценные свойства:

1. Обучение. Выбрав одну из моделей **ИНС**, создав сеть и выполнив алгоритм обучения, мы можем обучить **ИНС** решению задачи, которая ей по силам. Нет никаких гарантий, что это удастся сделать при выбранных: сети, алгоритме и задаче. Но, если все сделано правильно, то обучение бывает успешным.
2. Способность к обобщению. После обучения **ИНС** становится нечувствительной к малым изменениям входных сигналов (шуму или вариациям входных образов) и дает правильный результат на выходе.
3. Способность к абстрагированию. Если предъявить сети несколько искаженных вариантов входного образа, то **ИНС** сама может создать на выходе идеальный образ, с которым она никогда не встречалась.

Автору представляется, что использование **ИНС** на таком важном и обширном поле деятельности как определение термодинамических свойств веществ сильно отстает от требований современной науки и, в особенности, практики.

Фактически ни один практический расчет какого-либо теплообменного устройства не может обойтись без вычисления различных термодинамических (и физических тоже!) свойств сред (материалов, теплоносителей и так далее), использующихся в нем. При решении ряда важных практических задач расчет свойств занимает львиную долю в общем объеме вычислений.

Сюда следует отнести задачи оптимизации теплообменных устройств, расчеты аварийных и переходных процессов, имитационное моделирование и так далее.

Хорошо, если свойства среды, с которой Вы имеете дело, кем-то обобщены и описаны соответствующими уравнениями, легко поддающимися реализации на ЭВМ. А что делать, если в Вашем распоряжении - лишь таблица с экспериментально определенными (да и то не во всем необходимом Вам диапазоне) точками состояния?

Автор считает, что использование уникальных возможностей **ИНС** позволяет оптимально решать все вышеупомянутые проблемы:

Le **RNA** – c'est un moyen du traitement de l'information :

- a) C'est un modèle flexible pour l'approximation non linéaire des fonctions multidimensionnelles;
- b) C'est un moyen de la prévision dans le temps pour les processus qui dépendent de plusieurs variables;
- c) C'est un classificateur selon les plusieurs indices, qui fait la partition de l'espace d'entrée en domaines (zones, régions?); C'est un moyen de la perception des images;
- d) C'est un outil pour repérer selon les associations;
- e) C'est un modèle pour la recherche des régularités dans les paquets de données.

Les **RNA** sont libres des limitations (restrictions) des ordinateurs usuels grâce au traitement parallèle et à une forte liaison des neurones.

Les **RNA** modernes selon leurs complexité et "intelligence" s'approchent du système nerveux de l'insecte, et manifestent en ce moment les propriétés précieuses :

1. L'enseignement. Ayant choisi un des modèles du **RNA**, ayant créé le réseau et ayant accompli l'algorithme de l'enseignement, nous pouvons faire apprendre le **RNA** à la résolution d'un problème, qui soit au niveau de ses forces. Il n'y a pas de garanties que cette tâche va réussir si nous avons choisi : le réseau, l'algorithme et le problème (l'objectif). Mais, au cas où tout est fait correctement, il arrive que l'enseignement s'avère comme réussi.
2. La capacité de la généralisation. Après avoir passé par l'enseignement le **RNA** devient insensible aux petits changements des signaux d'entrée (le bruit ou les variations des images d'entrée) et donne le résultat juste à la sortie.
3. La capacité de l'abstraction. Si l'on présente au réseau quelques variantes déformées de l'image d'entrée, le **RNA** peut créer sur la sortie une image idéale, qu'il n'avait jamais rencontré.

L'auteur du présent article croit que l'utilisation des **RNA** pour un champ d'activité aussi vaste et important comme la définition des propriétés thermodynamiques des substances retarde fort des exigences de la science moderne et, surtout, de la pratique.

En réalité pas un seul calcul pratique d'un échangeur de chaleur ne peut pas se passer du calcul des diverses propriétés thermodynamiques (et physiques également!) des milieux (des matériaux, des fluides caloporteurs et cetera), appliquées dans celui-ci. Au cours de décision d'une série de tâches pratiques importantes le calcul des propriétés occupe la part du lion dans le volume total des calculs.

Ici il faut ranger les objectifs de l'optimisation des dispositifs - échangeurs de chaleur, les calculs des processus d'avarie et de transition, la simulation sur modèle et cetera.

C'est bien, si les propriétés du milieu que vous avez à étudier, sont généralisées par quelqu'un et sont décrites avec des équations correspondantes facilement programmables sur l'ordinateur. Mais qu'est-ce qu'il faut faire, au cas où vous n'avez dans votre disposition qu'un tableau avec les points de l'état expérimentalement définis (et ces définitions n'englobent pas toute la gamme dont vous avez besoin) ?

L'auteur trouve que l'utilisation des possibilités uniques des **RNA** permet de résoudre à l'optimum tous les objectifs susmentionnés :

1. - augmenter la vitesse opératoire du calcul des propriétés;
2. - obtenir des programmes finis du calcul des propriétés ayant pour base des données expérimentales;

1. - повышение быстродействия вычисления свойств;
2. - получение готовых программ расчета свойств на основе экспериментальных данных;
3. - обоснование распространения свойств веществ на области, в которых пока нет экспериментально подтвержденных данных.

Обоснование необходимости применения ИНС.

Теоретической предпосылкой к описанию термодинамических свойств веществ являются выводы термодинамики (см. [7]) из которых следует, что, как правило, свойства определяются по одному или двум термодинамическим параметрам. На практике этими параметрами в подавляющем большинстве являются давление P и температура T , т.к. эти параметры легко измеримы при экспериментальном определении свойств веществ.

В практических теплотехнических расчетах часто используются: энтальпия - h , удельный объем - v , внутренняя энергия - u .

Все эти, и многие другие термодинамические параметры связаны между собой уравнениями термодинамики, с помощью которых можно выразить одни пары параметров через другие пары параметров.

На практике задача вычисления термодинамических свойств связана с особенностью решаемой проблемы.

Например, при стационарных расчетах гидродинамики и теплопередачи удобно определять все свойства через P и h , а при нестационарных расчетах – через v и u .

Международные организации, такие, например, как *IAPWS* [8], не «забыли» о практических нуждах инженеров и исследователей, и автору не известно ни одного уравнения состояния с независимыми парами параметров (P, h) или (v, u).

Многие преодолели эту трудность, создав методики и программы пересчета исходных уравнений состояния с использованием итерационных процедур либо с непосредственной интерполяцией таблиц свойств.

Автор сам в 1975-1980 г. создал комплекс программ расчета всех термодинамических свойств воды и водяного пара по различным парам параметров на основе итерационных процедур и сплайн-интерполяции [9, 10].

Выход новых, более точных уравнений состояния во многом обесценивает сделанную ранее работу, т.к. с новыми данными приходится переделывать готовые и оптимизированные программы!

Часто в практических расчетах нет нужды использовать программы и методики, охватывающие весь возможный диапазон изменения параметров, но, в узком диапазоне изменения параметров программа должна работать быстро и точно!

Автор утверждает, что все указанные выше, да и многие другие проблемы, разрешимы с помощью *ИНС*!

3. - аргументировать распространение свойств веществ на области, в которых пока нет экспериментально подтвержденных данных.

L'argumentation de la nécessité de l'application des *RNA*.

Le pré-supposé théorique pour la description des propriétés thermodynamiques des substances sont les conclusions de la thermodynamique (voir [7]) desquelles il suit qu'en général, les propriétés sont définies selon un ou deux paramètres thermodynamiques. Pratiquement ces paramètres sont dans la plupart des cas: la pression P et la température T , car ceux derniers sont facilement mesurables pendant la définition expérimentale des propriétés des substances.

Dans les calculs pratiques de la technique de la chaleur sont utilisés souvent: l'enthalpie - h , le volume spécifique - v , l'énergie intérieure - u .

Tout ceux-ci, et plusieurs autres paramètres thermodynamiques sont liés entre eux par les équations de la thermodynamique, et avec leur aide on peut exprimer les paires des paramètres par l'intermédiaire des autres paires des paramètres.

Pratiquement la tâche du calcul des propriétés thermodynamiques est liée à la particularité du problème que nous avons à résoudre.

Par exemple, quand on fait des calculs stationnaires de l'hydrodynamique et de la transmission de la chaleur il est commode de définir toutes les propriétés par P et h , mais pour les calculs non stationnaires - par v et u .

Les organismes internationaux, tels, par exemple, comme *IAPWS* [8], n'ont pas «pris soin» des besoins pratiques des ingénieurs et des chercheurs, et l'auteur ne connaît pas une seule équation d'état ayant les paires indépendantes des paramètres (P, h) ou (v, u).

Plusieurs auteurs ont surmonté cette difficulté, ayant créé les méthodes et les programmes de la conversion des équations d'état initiales avec l'utilisation des procédures itératives ou avec l'interpolation directe des tableaux des propriétés.

L'auteur lui-même en 1975-1980 a créé l'ensemble des programmes du calcul de toutes les propriétés thermodynamiques de l'eau et de la vapeur d'eau selon les diverses paires des paramètres ayant pour base des procédures itératives et la spline-interpolation [9, 10].

La parution des nouvelles et plus exactes équations d'état dévalorise dans la plupart des cas le travail fait auparavant, car avec de nouvelles données il faut refaire les programmes finis et optimisés!

Souvent dans les calculs pratiques il n'y a pas de raison d'utiliser les programmes et les méthodes embrassant toute la gamme possible du changement des paramètres, mais, dans la gamme étroite du changement des paramètres le programme doit fonctionner vite et avec précision!

L'auteur affirme que tous les problèmes mentionnés ci-dessus, et plusieurs autres problèmes, peuvent être résolus avec l'aide du *RNA*!

Les définitions principales.

Nous examinerons la fonction multidimensionnelle $y = f(x)$, où le vecteur y a N_o de composantes, mais le vecteur x - N_I de composantes. Le moyen le plus simple de la formalisation - c'est d'utiliser le *RNA* avec N_I d'entrées et N_o de sorties.

Основные определения.

Рассмотрим многомерную функцию $y = f(x)$, где вектор y имеет N_o компонент, а вектор x — N_I компонент. Самый простой способ формализации — использовать **ИНС** с N_I входами и N_o выходами.

Компоненты вектора x подаются на вход сети, y — снимаются на выходе. **ИНС** обучается на известных значениях функции f .

Выбор количества нейронов и слоев.

Нет строго определенной процедуры для выбора количества нейронов и количества слоев в сети.

Чем больше количество нейронов и слоев, тем шире возможности сети, тем медленнее она обучается и работает, и тем более нелинейной может быть зависимость «вход-выход».

Количество нейронов и слоев связано:

1. со сложностью задачи;
2. с количеством данных для обучения;
3. с требуемым количеством входов и выходов сети;
4. с имеющимися ресурсами: памятью и быстродействием машины, на которой моделируется **ИНС**.

Были попытки записать эмпирические формулы для числа слоев и нейронов, но применимость формул оказалась очень ограниченной.

Если в сети слишком мало нейронов или слоев:

1. **ИНС** не обучится и ошибка при работе сети останется большой;
2. на выходе сети не будут передаваться резкие колебания аппроксимируемой функции $y(x)$.

Превышение требуемого количества нейронов тоже мешает работе сети.

Если нейронов или слоев слишком много:

1. быстродействие будет низким, а памяти потребуется много — на фон-неймановских ЭВМ;
2. **ИНС переобучится**: выходной вектор будет передавать незначительные и несущественные детали в изучаемой зависимости $y(x)$, например, шум или ошибочные данные;
3. зависимость выхода от входа окажется резко нелинейной: выходной вектор будет существенно и непредсказуемо меняться при малом изменении входного вектора x ;
4. **ИНС** будет неспособна к *обобщению*: в области, где нет или мало известных точек функции $y(x)$ выходной вектор будет случаен и непредсказуем, не будет адекватен решаемой задаче.

Подготовка входных и выходных данных

Данные, подаваемые на вход сети и снимаемые с выхода, должны быть правильно подготовлены.

Один из распространенных способов — нормализация (масштабирование):

$$X = (x - m) * C \quad 1$$

где x — исходный вектор;

X — нормализованный вектор;

Les composantes du vecteur x sont donnés pour l'entrée du réseau, celles du vecteur y - sont retirées à la sortie. Le **RNA** est enseigné avec les valeurs connues de la fonction f .

Le choix de la quantité de neurones et de couches.

Il n'y a pas de procédure strictement définie pour le choix de la quantité de neurones et de la quantité de couches dans le réseau.

Plus la quantité de neurones et de couches est importante, plus larges sont les possibilités du réseau, plus lentement ce dernier apprend et travaille, et la dépendance dite «entrée-sortie» pourrait être beaucoup plus non linéaire.

La quantité de neurones et de couches est liée :

1. Avec la complexité du problème;
2. Avec la quantité de données pour l'enseignement;
3. Avec la quantité demandée d'entrées et de sorties du réseau;
4. Avec les ressources disponibles : avec la mémoire et avec la vitesse opératoire de l'ordinateur, sur lequel on modèle le **RNA**.

Il y avait des tentatives d'enregistrer les formules empiriques pour le nombre de couches et de neurones, mais il s'est avéré que l'applicabilité de formules est fort limitée.

Au cas où dans le réseau il y ait très peu de neurones ou de couches :

1. Le **RNA** n'apprendra pas et la bogue au cours du fonctionnement du réseau va rester importante;
2. A la sortie du réseau les oscillations brusques de la fonction approximée $y(x)$ ne seront pas transmises.

L'excès de la quantité demandée de neurones empêche également le fonctionnement du réseau.

Au cas où la quantité de neurones ou de couches soit trop grande :

1. La vitesse opératoire serait basse, mais il faudrait beaucoup de mémoire - pour les ordinateurs dits «de Von-Neumann»;
2. Le **RNA** va *reapprendre* : le vecteur de sortie va transmettre les détails peu importants et insignifiants dans la dépendance étudiée $y(x)$, par exemple, le bruit ou des données erronées;
3. La dépendance de la sortie de l'entrée se trouvera brusquement non linéaire : le vecteur de sortie sera varié d'une manière importante est imprévisible avec un petit changement du vecteur d'entrée x ;
4. Le **RNA** sera incapable de *faire la généralisation* : dans le domaine, où il n'y a pas ou il y a très peu de points connus de la fonction $y(x)$ le vecteur de sortie serait accidentel et imprévisible, il ne sera pas adéquat au problème à résoudre.

La préparation des données d'entrée et celles de sortie

Les éléments données pour l'entrée du réseau et retirées de la sortie, doivent être correctement préparées.

Un des moyens de ceux qui sont répandus - la normalisation (mise à l'échelle) :

$$X = (x - m) * C \quad 1$$

Où le x - c'est le vecteur initial;

Le X - c'est le vecteur normalisé;

Le vecteur m - c'est la valeur moyenne de l'ensemble des entrées;

Le C - c'est le facteur d'échelle.

Вектор m — усредненное значение совокупности входных данных;
 C — масштабный коэффициент.

Нормализация желательна, чтобы привести данные в допустимый диапазон. Если этого не сделать, то возможно несколько проблем:

1. нейроны входного слоя или окажутся в постоянном насыщении ($|m|$ велико, дисперсия входных данных мала), или будут все время заторможены ($|m|$ мало, дисперсия мала);
2. весовые коэффициенты примут очень большие или очень малые значения при обучении (в зависимости от дисперсии), и, как следствие, растянется процесс обучения и снизится точность.

Определим методы обучения и критерии качества обучения нейросети.

Алгоритмы обучения бывают с учителем и без. Алгоритм называется *алгоритмом с учителем*, если при обучении известны и входные, и выходные векторы сети. Имеются пары «вход + выход» — «известные условия задачи и решение». В процессе обучения **ИНС** меняет свои параметры и учится давать нужное отображение $X \rightarrow Y$. **ИНС** учится давать результаты, которые нам уже известны. За счет способности к обобщению сетью могут быть получены новые результаты, если подать на вход вектор, который не встречался при обучении.

Алгоритм относится к обучению *без учителя*, если известны только входные векторы, и на их основе **ИНС** учится давать *наилучшие* значения выходов. Что понимается под «*наилучшими*» — определяется алгоритмом обучения.

Перцептрон обучается с учителем. Это означает, что должно быть задано множество пар векторов:

$\{x^n, d^n\}$, $n = \{1, \dots, N\}$, где $\{x^n\} = \{x^1, \dots, x^N\}$ — формализованное условие задачи, а $\{d^n\} = \{d^1, \dots, d^N\}$ — известное решение (истинный выход) для этого условия.

Совокупность пар $\{x^n, d^n\}$ составляет *обучающее множество*.

Количество элементов в обучающем множестве N — должно быть достаточным для обучения сети, чтобы под управлением алгоритма сформировать набор параметров сети, дающий нужное отображение $X \rightarrow Y$.

Количество пар в обучающем множестве не регламентируется. Если элементов слишком много или мало, **ИНС** не обучится и не решит поставленную задачу.

Выберем один из векторов x^s и подадим его на вход сети. На выходе получится некоторый вектор y^s . Тогда ошибкой сети можно считать $E = \|d^s - y^s\|$ для каждой пары $\{x^n, d^n\}$.

Задача обучения перцептрона ставится так: подобрать такие значения параметров сети, чтобы ошибка была минимальна для данного обучающего множества $\{x^n, d^n\}$.

Большая часть методов обучения — итерационные. Параметрам сети (весовым коэффициентам и пороговым уровням) присваиваются малые начальные значения. Затем параметры изменяются так, чтобы ошибка E убывала. Изменения продолжаются до тех пор, пока ошибка не станет достаточно малой.

La normalisation est désirable afin d'amener les données à la gamme admissible. Au cas où l'on ne fait pas cela, quelques problèmes vont surgir :

1. Les neurones de la couche d'entrée vont se trouver dans la saturation constante ($|m|$ est grand, la dispersion des données d'entrée est petite), ou bien tout le temps ils seront inhibés ($|m|$ est petit, la dispersion est petite);
2. Les constantes de poids vont prendre de très grandes ou très petites valeurs au cours de l'enseignement (en fonction de la dispersion), et, comme suite, le procès de l'enseignement va devenir plus long et la précision va baisser.

Nous avons à définir les méthodes de l'enseignement et les critères de la qualité de l'enseignement du réseau neuronique.

Il existe les algorithmes de l'enseignement dits « avec le professeur » et « sans le professeur ». L'algorithme dit « *l'algorithme avec le professeur* » peut être appelé ainsi au cas où au cours de l'enseignement sont connus les vecteurs d'entrée comme les vecteurs de sortie du réseau. Il y a des paires «l'entrée + la sortie» - «les données connues d'un problème et la résolution». Au cours de l'enseignement le **RNA** change ses paramètres et apprend à donner l'application (la représentation) nécessaire $X \rightarrow Y$. Le **RNA** apprend à donner les résultats, qui nous sont déjà connus. Avec la faculté de la généralisation le réseau peut obtenir de nouveaux résultats, au cas où à l'entrée on donne le vecteur, que l'on n'a pas encore rencontré au cours de l'enseignement.

L'algorithme peut être rapporté à l'enseignement dit « *sans le professeur* », si seuls les vecteurs d'entrée nous sont connus, et sur leur base le **RNA** apprend à donner les valeurs *les meilleurs* des sorties. Qu'est-ce qui est compris sous les "meilleurs" - est défini par l'algorithme de l'enseignement.

Le perceptrone apprend avec le professeur. Cela veut dire que toute une multitude de paires de vecteurs doit être donnée:

$\{x^n, d^n\}$, $n = \{1 \dots N\}$, où $\{x^n\} = \{x^1, \dots, x^N\}$ - c'est la donnée formalisée du problème, et $\{d^n\} = \{d^1, \dots, d^N\}$ - c'est la résolution connue (la sortie véritable) pour cette donnée.

La totalité des paires $\{x^n, d^n\}$ fait *l'ensemble enseignant*.

La quantité d'éléments dans l'ensemble enseignant N - doit être suffisante pour l'enseignement du réseau pour que sous la direction de l'algorithme nous puissions former un assortiment des paramètres du réseau, donnant l'application nécessaire $X \rightarrow Y$.

La quantité de paires dans la multitude apprenant n'est pas réglementée. Au cas où il y a trop, ou peu d'éléments, le **RNA** n'apprendra pas et ne résoudra pas le problème en question.

Nous allons choisir un des vecteurs x^s et nous le donnerons à l'entrée du réseau. Sur la sortie nous obtiendrons un certain vecteur y^s . Alors nous pourrons estimer comme la bogue du réseau $E = \|d^s - y^s\|$ pour chaque paire $\{x^n, d^n\}$.

L'objectif de l'enseignement du perceptrone est formulé ainsi : choisir de telles valeurs des paramètres du réseau afin que la bogue soit minimale pour cet ensemble enseignant $\{x^n, d^n\}$.

La grande partie des méthodes de l'enseignement – sont les méthodes itératifs. Aux paramètres du réseau (les constantes de poids et les niveaux de seuil) on attribue de petites valeurs initiales. Puis les paramètres changent de telle manière, pour que la bogue E décroisse. Les changements continuent jusqu'à ce que la bogue ne deviendra pas assez petite.

L'estimation la plus simple E – dite la bogue de pourcentage, on la définit comme :

Наиболее простая оценка E - т.н. процентная ошибка, определяемая как:

$$\%Error = \frac{100\%}{N * P} * \sum_{j=0}^P \sum_{i=0}^N \frac{|dy_{ij} - dd_{ij}|}{dd_{ij}}, \quad 2$$

где dy_{ij} - денормализованный выход сети для *i-того* экземпляра данных в *j-том* выходном элементе сети;

dd_{ij} - денормализованный истинный выход для *i-того* экземпляра данных в *j-том* выходном элементе сети.

Заметим, что $\%Error$ может легко вводить в заблуждение. Например, пусть ваши выходные данные находятся в диапазоне от 0 до 100. Для одного образца ваш истинный выход - 0.1, и ваш фактический выход - 0.2. Даже при том, что две величины весьма близки, $\%Error$ для этого образца = 100!

Среднеквадратическая погрешность (MSE) может являться отличным критерием (к сожалению, не единственным!) успеха нейросетевой аппроксимации:

$$MSE = \frac{\sum_{j=0}^P \sum_{i=0}^N (d_{ij} - y_{ij})^2}{N * P}, \quad 3$$

где:

P - число выходных элементов сети,

N - количество элементов в обучающем множестве.

Вторым критерием погрешности для учета «веса» каждой переменной может служить **нормализованная среднеквадратическая погрешность (NMSE)**:

$$NMSE = \frac{P * N * MSE}{\sum_{j=1}^P \frac{N * \sum_{i=0}^N d_{ij}^2 - (\sum_{i=0}^N d_{ij})^2}{N}}, \quad 4$$

Мы нашли способ вычислить уравнение регресса, но мы все еще не имеем меры того, как успешно кривая регресса представляет отношения между x и d . Величина среднеквадратической ошибки (MSE) может лишь определить, какая кривая лучше всего соответствует данным, но не обязательно доказывает, хорошо ли соответствует она данным, потому что MSE зависит от абсолютной величины образцов данных. Например, просто изменяя масштабы данных, мы можем изменить MSE , не изменяя самих данных. **Коэффициент корреляции** решает эту проблему.

Коэффициент корреляции (r) между двумя случайными переменными x и d определяется как:

$$r = \frac{\sum_{i=1}^N (x_i - \bar{x}) \times (d_i - \bar{d})}{\sqrt{\frac{\sum_{i=1}^N (d_i - \bar{d})^2}{N}} \times \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N}}}, \quad 5$$

$$\%Error = \frac{100\%}{N * P} * \sum_{j=0}^P \sum_{i=0}^N \frac{|dy_{ij} - dd_{ij}|}{dd_{ij}}, \quad 2$$

où dy_{ij} - est la sortie dénormalisée du réseau pour *i-nième* exemplaire des données dans *j-ième* élément de sortie du réseau;

dd_{ij} - est la sortie dénormalisée véritable pour *i-nième* exemplaire des données dans *j-ième* élément de sortie du réseau.

Il est à remarquer que $\%Error$ peut facilement induire en erreur. Par exemple, supposons que vos données de sortie se trouvent dans la gamme de 0 jusqu'à 100. Pour un modèle votre sortie véritable - est 0.1, et votre sortie réelle - est 0.2. Même si les deux valeurs sont très proches, $\%Error$ pour ce modèle = 100!

L'erreur moyenne quadratique (MSE) peut être un critère excellent (malheureusement, pas l'unique!) du succès de l'approximation avec l'aide du **RNA** :

$$MSE = \frac{\sum_{j=0}^P \sum_{i=0}^N (d_{ij} - y_{ij})^2}{N * P}, \quad 3$$

où :

P - est le nombre des éléments de sortie du réseau,

N - est la quantité d'éléments dans l'ensemble enseignant.

L'erreur normalisée moyenne quadratique (NMSE) peut servir du deuxième critère de l'erreur pour le compte du "poids" de chaque variable :

$$NMSE = \frac{P * N * MSE}{\sum_{j=1}^P \frac{N * \sum_{i=0}^N d_{ij}^2 - (\sum_{i=0}^N d_{ij})^2}{N}}, \quad 4$$

Nous avons trouvé le moyen de calculer l'équation de la régression, mais nous n'avons pas encore de mesure pour estimer, avec quel degré de réussite la courbe de la régression représente les relations entre x et d . La valeur de la bogue moyenne quadratique (MSE) ne peut que définir, quelle courbe correspond le mieux aux données, mais ne prouve absolument, si elle correspond bien aux données, parce que la MSE dépend de la valeur absolue des modèles des données. Par exemple, en changeant tout simplement les échelles des données, nous pouvons changer la MSE , sans changer les données elles-mêmes. **Le taux de la corrélation** résout ce problème.

Le taux de la corrélation (r) entre les deux variable accidentelles x et d est défini comme :

$$r = \frac{\sum_{i=1}^N (x_i - \bar{x}) \times (d_i - \bar{d})}{\sqrt{\frac{\sum_{i=1}^N (d_i - \bar{d})^2}{N}} \times \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N}}}, \quad 5$$

$$\text{где: } \bar{d} = \frac{\sum_{i=1}^N d_i}{N}; \quad \bar{x} = \frac{\sum_{i=1}^N x_i}{N}.$$

Если r является неположительной величиной, то никакой корреляции между входными и выходными данными нет. Если r находится в диапазоне от 0 до 1, то наблюдается корреляция, и чем ближе r к 1, тем эта корреляция лучше.

Информационный критерий (AIC) используется, чтобы измерить компромисс между эффективностью обучения и размером сети. Цель состоит в том, чтобы минимизировать этот критерий, и произвести ИНС с наилучшим обобщением:

$$AIC(k) = N * \ln(MSE) + 2 * k, \quad 6$$

где: k – число весовых коэффициентов, произведенных нейросетью.

Минимальная длина описания Риссейнна (MDL), критерий, подобный AIC, в котором комбинируется ошибка модели и количество степеней свободы для определения уровня обобщения. Цель состоит в том, чтобы минимизировать этот критерий:

$$MDL(k) = N * \ln(MSE) + \frac{1}{2} * k * \ln(N) \quad 7$$

Итак, определив критерии обучения, можем перейти к самому обучению.

Общая схема обучения перцептрона:

1. Инициализировать веса и параметры функции активации в малые ненулевые значения;
2. Подать на вход один образ и рассчитать выход;
3. Посчитать ошибку E , сравнив d и y .
4. Изменить веса и параметры функции активации так, чтобы ошибка E уменьшилась.
5. Повторить шаги 2-4 до тех пор, пока ошибка не перестанет убывать или не станет достаточно малой.

Здесь веса меняются так, что убывает не E , а E^s , относящаяся к образу s , а не ко всему обучающему множеству. Шаги в данном варианте алгоритма делаются не в направлении убывания E , а в направлении убывания E^s , таким образом, E не обязательно должна убывать. Какие условия необходимы для существенной убывания E ? Опыт показывает, что для этого необходимо отсутствие упорядоченности в предъявлении образов, т.е. в выборе s на каждой итерации. Если образы выбираются случайно из обучающего множества, то ошибка E чаще всего убывает. Если же есть упорядоченность (например, образы предъявляются циклически: 1-й, 2-й, ..., S -й, 1-й, ...) то чаще всего $E(t)$, где t — время обучения, не имеет предела при $t \rightarrow \infty$ т.е. алгоритм расходится. В этом случае E^s тоже убывает при каждом изменении параметров, но при следующей коррекции для образа $(s + 1)$ ошибка E^{s+1} убывает, а E^s , относящаяся к предыдущему образу, возрастает сильнее, так что E может увеличиться. ИНС "забывает" текущий образ при предъявлении следующего.

Чтобы шаг по параметрам на каждой итерации делался в правильном направлении, надо провести усреднение по S , т.е. предъявить все образы, и коррекции вычислять по всем образам сразу. Такие алгоритмы называются

$$\text{où: } \bar{d} = \frac{\sum_{i=1}^N d_i}{N}; \quad \bar{x} = \frac{\sum_{i=1}^N x_i}{N}.$$

Si r est la valeur non positive, il n'y ait aucune corrélation entre les données d'entrée et celles de la sortie. Si r se trouve dans la gamme de 0 jusqu'à 1, on observe la corrélation, et plus r est près de 1, plus cette corrélation soit meilleure.

Le critère d'information (AIC) est utilisé pour mesurer le compromis entre l'efficacité de l'enseignement et la taille du réseau. L'objectif vise à minimiser ce critère, et à produire (créer) le RNA avec une généralisation la meilleure possible:

$$AIC(k) = N * \ln(MSE) + 2 * k, \quad 6$$

où: k - le nombre des constantes de poids, produites par le RNA.

La longueur minimale de la description de Rissanen (MDL), le critère qui est semblable à AIC, dans lequel on combine la bogue du modèle et la quantité de degrés de la liberté pour définir le niveau de la généralisation. Le but est de minimiser ce critère :

$$MDL(k) = N * \ln(MSE) + \frac{1}{2} * k * \ln(N) \quad 7$$

Donc, ayant défini les critères de l'enseignement, nous pouvons passer à l'enseignement lui-même.

Le schéma général de l'enseignement du perceptrone :

1. Initialiser les poids et les paramètres de la fonction de l'activation à (vers) de petites valeurs non nulles;
2. Donner pour l'entrée une image et calculer la sortie;
3. Calculer la bogue E , ayant comparé d et y .
4. Changer les poids et les paramètres de la fonction de l'activation de manière que la bogue E diminue.
5. Répéter les actes 2-4 jusqu'à ce que la bogue ne cessera pas de diminuer ou ne deviendra pas assez petite.

Ici les poids varient de façon que ce n'est pas E qui diminue, mais E^s , cette dernière est en rapport avec l'image s , mais pas avec tout l'ensemble enseignant. Les actes dans la variante donnée de l'algorithme ne se font pas en direction de diminuer E , mais en direction de la perte de E^s , de cette façon, E ne doit diminuer absolument pas. Quelles conditions sont-elles nécessaires pour la perte sensible de E ? L'expérience nous montre que pour cela nous avons besoin de l'absence de l'ordre dans la présentation des images, autrement dit dans le choix de s sur chaque itération. Au cas où les images sont choisies par hasard de l'ensemble enseignant, la bogue E diminue le plus souvent. S'il y a un ordre (par exemple, les images sont présentées d'une manière cyclique : 1-er, 2-ème..., S -ème, 1-er...), le plus souvent $E(t)$, où t – est le temps de l'enseignement, n'a pas de limite avec $t \rightarrow \infty$ c'est-à-dire l'algorithme diverge. Dans ce cas E^s diminue aussi avec chaque changement des paramètres, mais à la correction suivante pour l'image $(s + 1)$ la bogue E^{s+1} diminue, mais E^s , qui se rapporte à l'image précédente, augmente plus fort, de sorte qu' E peut augmenter. Le RNA "oublie" l'image courante avec la présentation de celle qui suit.

Pour qu'un acte pour les paramètres sur chaque itération se fasse dans la direction juste, il faut faire le moyennage selon S , c'est-à-dire présenter toutes les images, et calculer les corrections selon toutes les images à la fois. Les susdits algorithmes s'appellent *les algorithmes*

алгоритмами с пакетной коррекцией (batch update). Они требуют больших затрат вычислительного времени и памяти, но сходятся за меньшее число итераций.

В большинстве случаев **E** при таком методе обучения сходится и достигает локального минимума. Для каждой конкретной задачи нет гарантий, что **E** сойдется к приемлемому значению за конечное число шагов.

Для однослойного перцептрона алгоритм обучения очевиден. Как обобщить этот простой алгоритм на случай многослойной сети? Эту задачу решает алгоритм Румельхарта-Хинтона-Вильямса (Rumelhart, Hinton, & Williams) (алгоритм обратного распространения ошибки). Он был предложен в различных вариациях в нескольких научных работах, существует также множество улучшенных версий алгоритма. Достаточно подробно этот алгоритм изложен в [3-5], и здесь нет необходимости его описывать.

Результаты применения **ИНС** для аппроксимации термодинамических свойств.

Наиболее изученным веществом с точки зрения термодинамических свойств является вода. Это и понятно, т.к. с использованием воды в качестве теплоносителя или рабочего тела работает подавляющее большинство теплопередающих устройств, машин и механизмов. Кроме того, в силу значительной нелинейности свойств воды, уравнения, описывающие ее термодинамические свойства, весьма сложны [8, 10] (область описания свойств разбита на несколько подобластей, в каждой из которых сами свойства описываются полиномами вплоть до 40 степени включительно!)

Автор выбрал воду и водяной пар еще и потому, что мог контролировать все результаты, получаемые **ИНС**, с помощью пакета программ **H2O** [9-10], и использовать этот пакет для «обучения» **ИНС**.

Из большого числа иерархий **ИНС** автор использовал для экспериментов «**ИНС** прямого распространения» с использованием многослойного перцептрона (**ИНС** МП) с функцией активации - гиперболическим тангенсом (здесь и далее используется устоявшаяся к настоящему времени терминология **ИНС**, см., например [1-4]) как наиболее простую и наглядную. Проводились эксперименты и с другими **ИНС**:

- Модулярные **ИНС**;
- Элмана – Джордана (Elman-Jordan) **ИНС**;
- **RBF**(*Radial Basis Function*) **ИНС**;
- *Рекуррентные* **ИНС** с полной и частичной рекурсией.

Однако каких либо преимуществ в исследуемом вопросе они не показали.

В практическом плане всем интересно применение единых алгоритмов для:

1. Расчета термодинамических свойств на линии насыщения;
2. Расчета термодинамических свойств во всей области состояния;

Проведенные вычислительные эксперименты не ставили целью получить оптимальную **ИНС**, однако показали следующие впечатляющие результаты:

1. Была построена **ИНС** для аппроксимации термодинамических свойств на линии

avec la correction de paquet (batch update). Ils demandent de grandes dépenses du temps calculatoire et de la mémoire, mais convergent avec une quantité plus petite des itérations.

Dans la plupart des cas **E** avec une telle méthode de l'enseignement converge et atteint un minimum local. Pour chaque problème concret il n'y a pas de garanties *qu'E* sera convergé vers la valeur acceptable pour (?) le nombre final des actes.

Pour un perceptrone à un couche l'algorithme de l'enseignement est évident. Comment généraliser ce simple algorithme pour le cas du réseau multicouche ? Cet objectif est résolu par l'algorithme de Rumelhart, Hinton, & Williams, (l'algorithme de la propagation inverse d'erreur). Il avait été proposé dans des versions diverses dans quelques travaux scientifiques, il existe également une multitude de versions améliorées de l'algorithme. Avec un certain nombre de détails cet algorithme est exposé dans [3-5], ici il n'y a pas de nécessité de le décrire.

Les résultats de l'application du **RNA** pour l'approximation des propriétés thermodynamiques.

La substance la plus étudiée du point de vue des propriétés thermodynamiques est l'eau. C'est bien clair, car avec l'utilisation de l'eau à titre du fluide caloporteur ou comme l'agent moteur fonctionne la majorité écrasante des dispositifs de transmission thermique, des machines et des mécanismes. En outre, à cause de la non linéarité considérable des propriétés de l'eau, les équations qui décrivent ses propriétés thermodynamiques, sont très complexes [8, 10] (le domaine de la description des propriétés est divisé en quelques sous domaines, et dans chacun de ceux derniers les propriétés elles-mêmes sont décrites par les polynômes jusqu'à 40-ème degrés y compris!).

L'auteur a choisi l'eau et la vapeur d'eau pour une raison de plus: c'est parce qu'il pouvait contrôler tous les résultats reçus par le **RNA** avec l'aide du progiciel **H2O** [9-10], et utiliser ce paquet pour l'"enseignement" du **RNA**.

D'un grand nombre des hiérarchies des **RNA** l'auteur utilisait pour les expériences «le **RNA** de la propagation directe» avec l'utilisation du perceptrone multicouche (**RNA** MC) avec la fonction de l'activation - la tangente hyperbolique (ici et ensuite on utilise la terminologie des **RNA** qui s'est formée vers le présent, voir, par exemple [1-4]) comme la plus simple et éclairante. On avait fait les expériences également avec d'autres **RNA** :

- Avec les **RNA** *Modulaires*;
- Avec les **RNA** d'Elman-Jordan;
- Avec les **RNA** *RBF* (*Radial Basis Function*);
- Avec les **RNA** *Récurrents* ayant la récursion complète et partielle.

Cependant ils n'ont pas montré des avantages dans la question étudiée.

Dans le plan pratique l'application des algorithmes universels est intéressante pour :

1. Le calcul des propriétés thermodynamiques à la ligne de la saturation;
2. Le calcul des propriétés thermodynamiques dans tout le domaine de l'état;

Les expériences calculatoires qui ont été faites ne mettaient pas pour but d'obtenir le **RNA** d'optimum, mais cependant ont montré les résultats impressionnants qui suivent :

1. On a construit le **RNA** pour l'approximation des propriétés thermodynamiques à la ligne de la saturation selon (d'après) la température de la saturation dans la gamme des températures de 273.16K jusqu'à 647.14K.

насыщения по температуре насыщения в диапазоне температур от 273.16К до 647.14К.

Вектор выходов *ИНС* включал в себя следующие величины: давление, производную давления по температуре, энтальпии, плотности, энтропии, изобарные теплоемкости, скорости звука, удельные вязкости, удельные теплопроводности и коэффициент поверхностного натяжения – всего 17 параметров.

Ниже приведены результаты двух экспериментов с одним и двумя скрытыми слоями нейронов (по 17 нейронов в каждом). Критерием окончания обучения *ИНС* принималось состояние, когда *ИНС* начала переобучаться. Переобучение – это процесс, когда *ИНС* просто запоминает обучающий набор и неспособна обобщить проблему.

На рис.1 изображены две таблицы из результатов расчетов по программе обучения *ИНС* и один график. Верхняя таблица представляет критерии, вычисленные по формулам (1-6) для *ИНС* с одним скрытым слоем нейронов, а нижняя таблица – то же для *ИНС* с двумя скрытыми слоями. На каждом цикле обучения (в терминологии *ИНС* – эпохе) проводилась проверка работы *ИНС* на независимом наборе (CV - Cross Validation) данных для определения правильности процесса обучения. На графике представлена типичная кривая изменения *MSE* в зависимости от номера эпохи обучения (верхняя кривая – обучающий набор, нижняя – CV).

Сам набор обучения составлял 80 векторов данных, набор CV – 25 векторов.

Из полученных данных видно, что *ИНС* с одним скрытым слоем нейронов уже дает прекрасный результат!

Сходимость, судя по кривой обучения Рис.1, достаточно хорошая, т.е. *ИНС* обучается за небольшое число эпох.

Le vecteur des sorties du *RNA* comprenait les valeurs suivantes : la pression, la dérivée de la pression selon la température, les enthalpies, les densités, les entropies, les capacités thermiques isobariques, les vitesses du son, les spécifiques de la viscosité, la spécifique de la conductibilité de la chaleur et l'énergie superficielle spécifique – au total les 17 paramètres.

Ci-dessous nous donnons les résultats de deux expériences avec un et avec deux couches latentes des neurones (les 17 neurones dans chacune). Nous considérons comme critère de la fin de l'enseignement du *RNA* l'état, quand le *RNA* commençait à reapprendre. Le re-enseignement – c'est le procès, quand le *RNA* retient simplement le paquet de l'enseignement et il est incapable de généraliser le problème.

Sur fig. 1 on représente deux tableaux faits des résultats des calculs selon le programme de l'enseignement du *RNA* et un graphique (courbe). Le tableau supérieur présente les critères calculés selon les formules (1-6) pour les *RNA* avec une couche latente des neurones, mais le tableau inférieur – représente la même situation pour le *RNA* ayant les deux couches latentes. Sur chaque cycle de l'enseignement (dans la terminologie des *RNA* le cycle – c'est l'époque) on faisait la vérification du travail du *RNA* avec un paquet indépendant des données (CV - Cross Validation) pour la définition de la justesse du procès de l'enseignement. Sur le graphique on présente la courbe typique du changement de *MSE* en fonction du numéro de l'époque de l'enseignement (la courbe supérieure - le paquet de l'enseignement, la courbe inférieure - CV).

Le paquet de l'enseignement faisait les 80 vecteurs des données, le paquet CV - les 25 vecteurs.

Des données reçues on voit que le *RNA* ayant une couche latente des neurones donne déjà un beau résultat!

La convergence, selon la courbe de l'enseignement de Fig. 1, est assez bonne, c'est-à-dire le *RNA* apprend pendant une quantité pas très grande des époques.

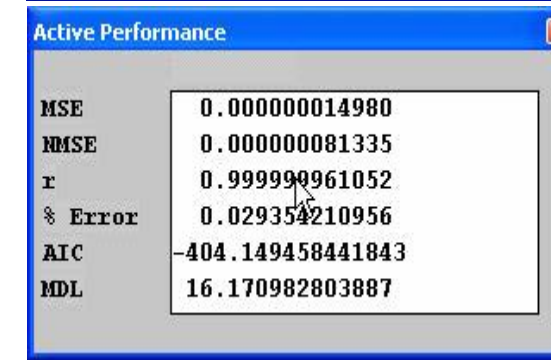
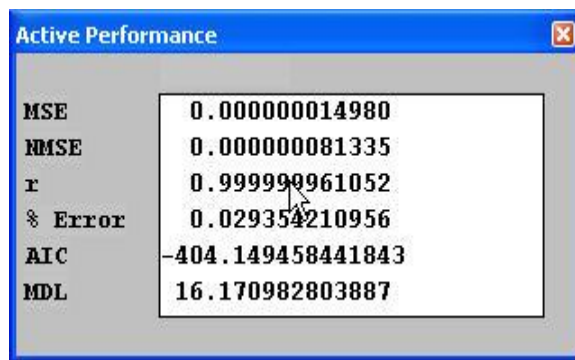
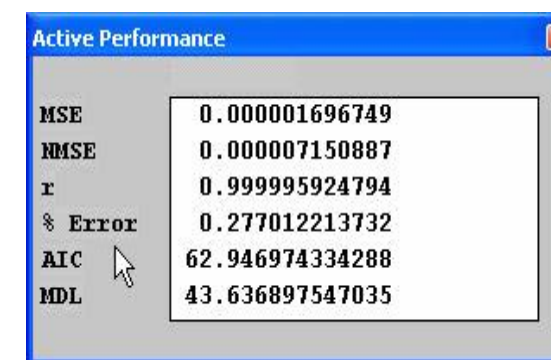
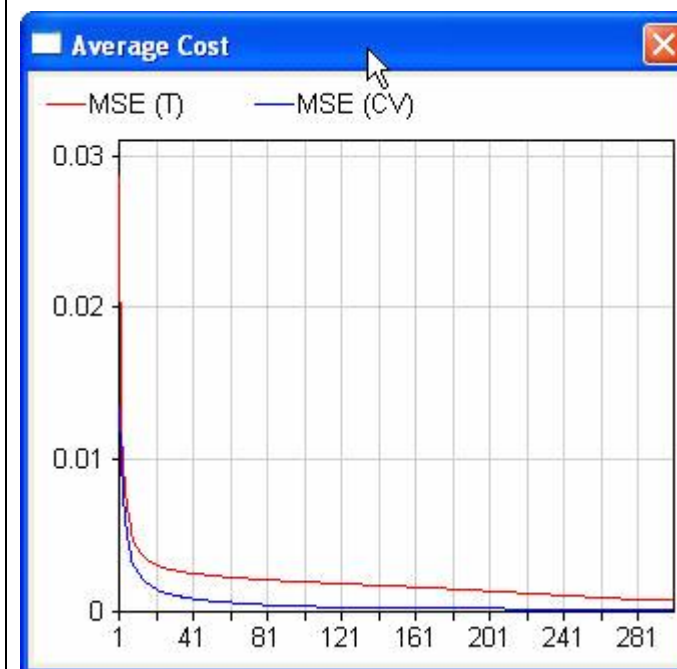
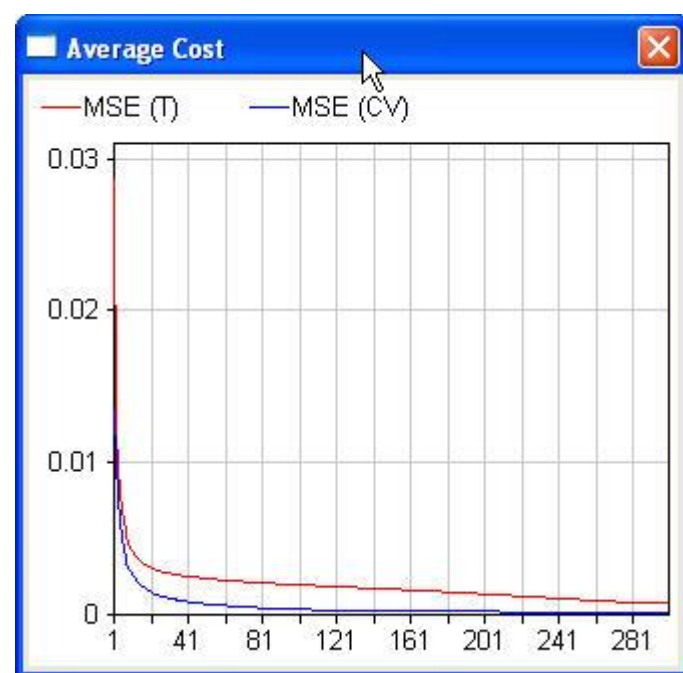
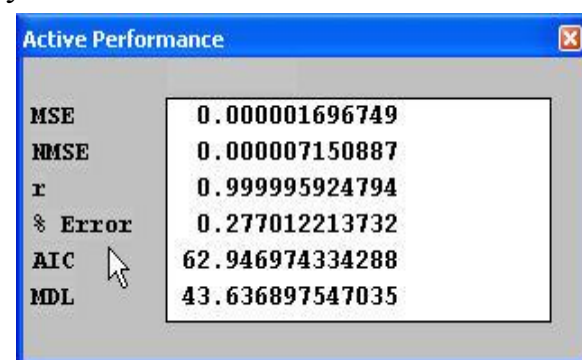


fig. 1

Рис. 1

2. Была построена *ИНС* для аппроксимации термодинамических свойств во всей области состояния по температуре в диапазоне от 273.16К до 647.14К и плотности в диапазоне от $1004 \frac{\text{кг}}{\text{м}^3}$, до $0,4 \frac{\text{кг}}{\text{м}^3}$.

Вектор выходов *ИНС* включал в себя следующие величины: давление, энтальпия и ее производные по температуре и плотности, энтропия и ее производные по температуре и плотности, внутренняя энергия и ее производные по температуре и плотности, изохорная теплоемкость, скорость звука, функция Гельмгольца и ее производные по температуре и плотности, функция Гиббса – всего 18 параметров.

Ниже приведены результаты двух экспериментов с одним и двумя скрытыми слоями нейронов (по 18 нейронов в каждом). Критерием окончания обучения *ИНС* принималось состояние, когда *ИНС* начинала переобучаться.

На рис.2 изображены две таблицы из результатов расчетов по программе обучения *ИНС* для всей области свойств воды и пара. Левая таблица представляет критерии, вычисленные по формулам (1-6) для *ИНС* с одним скрытым слоем нейронов, а правая таблица - то же для *ИНС* с двумя скрытыми слоями. Сам набор обучения составлял 1400 векторов данных, набор CV – 600 векторов.

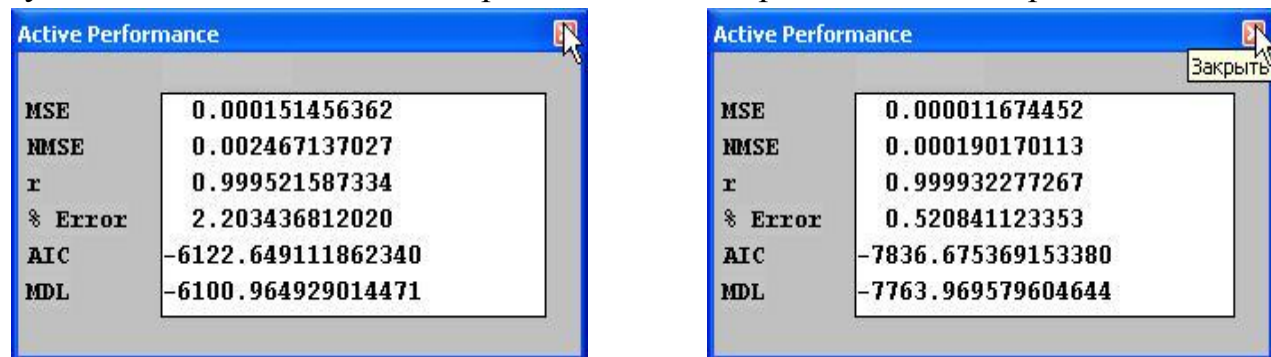


Рис. 2

Проведенные эксперименты с областями, включающими в себя метастабильные состояния воды и пара, заключенные между спинодалью и бинодалью дают не менее высокоточные результаты.

Выводы

1. *ИНС* прекрасно проводит аппроксимацию (интерполяцию) сложной функции.
2. В отличие от интерполяционного полинома *ИНС* хорошо предсказывает значение функции за пределами области определения.
3. Получаемая структура *ИНС* прекрасно подходит для практических расчетов, т.к. получается подпрограмма единообразного вычисления,

2. On a construit le *RNA* pour faire l'approximation des propriétés thermodynamiques dans tout le domaine de l'état selon la température dans la gamme de 273.16K jusqu'à 647.14K et selon la

densité dans la gamme de $1004 \frac{\text{кг}}{\text{м}^3}$, jusqu'à $0,4 \frac{\text{кг}}{\text{м}^3}$.

Le vecteur des sorties du *RNA* comprenait les valeurs suivantes : la pression, l'enthalpie et ses dérivées selon la température et la densité, l'entropie et ses dérivées selon la température et la densité, l'énergie intérieure et ses dérivées selon la température et la densité, la capacité thermique isochore, la vitesse du son, la fonction de Helmholtz et ses dérivées selon la température et la densité, la fonction de Gibbs – au total les 18 paramètres.

Plus bas on présente les résultats de deux expériences avec un et avec deux couches latentes des neurones (les 18 neurones dans chacune). Comme critère de la fin de l'enseignement du *RNA* on acceptait l'état, quand le *RNA* commençait à reapprendre.

Sur fig. 2 sont représentés deux tableaux des résultats des calculs d'après le programme de l'enseignement du *RNA* pour tout le domaine des propriétés de l'eau et de la vapeur. Le tableau gauche présente les critères calculés selon les formules (1-6) pour le *RNA* avec une couche latente des neurones, mais le tableau droit – est la même chose pour le *RNA* avec deux couches latentes. L'ensemble de l'enseignement faisait 1400 de vecteurs des données, l'ensemble (le paquet) CV - comprenait 600 vecteurs.

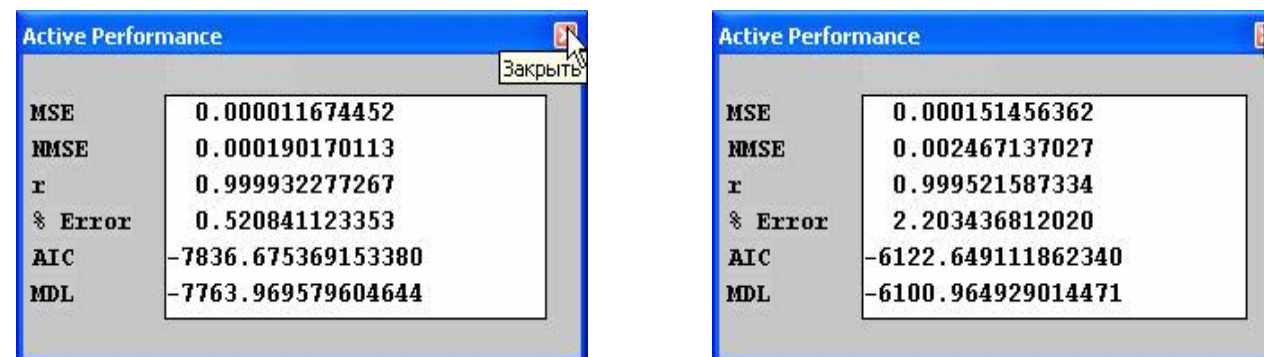


Fig. 2

Les expériences faites avec les domaines comprenant les états métastables de l'eau et de la vapeur, qui se trouvent entre la courbe de limite de l'état métastable (spinodale) et la courbe de saturation (binodale) donnent des résultats qui ne sont pas moins précis.

Les conclusions

1. Le *RNA* fait parfaitement l'approximation (interpolation) de la fonction complexe.
2. À la différence du polynôme d'interpolation le *RNA* prédit bien la valeur de la fonction en dehors du domaine de la définition.
3. La structure obtenue du *RNA* convient parfaitement pour les calculs pratiques, car on arrive à un sous-routine du calcul uniforme qui est bien réalisé sur les processeurs avec grand Cash ou sur les ordinateurs parallèles. La différence contrôlée entre la vitesse de fonctionnement des programmes réalisant les *RNA* et entre les dépendances ordinaires sur le processeur AMD 64-3500 - a fait 18 fois!
4. Les auteurs ne mettaient pas comme objectif de ce travail l'obtention des résultats les meilleurs, bien qu'il est évident qu'avec la sélection spéciale de la structure du *RNA* et des fonctions de l'activation des neurones dans le perceptrone, on peut obtenir les

хорошо реализуемая на процессорах с большим КЭШем (cash) или на параллельных компьютерах. Проверенная разница между быстродействием программ, реализующих ИНС и между обычными зависимостями на процессоре AMD 64-3500 - составила 18 раз!

4. Авторы не ставили целью этой работы получение наилучших результатов, хотя очевидно, что при специальном подборе структуры ИНС и функций активации нейронов в перцептроне, можно добиться значительно лучших показателей точности и быстродействия.
5. Походы, использованные авторами при аппроксимации термодинамических свойств, могут быть обобщены и на другое константное обеспечение расчетов на ЭВМ тепловых, гидродинамических и физико-химических процессов. Это актуально, например, при создании нового поколения расчетных кодов для анализа аварийных и переходных процессов в ЯЭУ (Ядерных Энергетических Установках).

Список литературы

1. Колмогоров А.Н. «О представлении непрерывных функций нескольких переменных суперпозициями непрерывных функций меньшего числа переменных». Докл. АН СССР, 1956. Т. 108, №. 2 С.179-182.
2. Горбань А.Н. «Обобщенная аппроксимационная теорема и вычислительные возможности нейронных сетей». Вычислительный центр СО РАН в г.Красноярске, 1993.
3. Суворцев И.С., Клюкин В.И., Пивоварова Р.П. «Нейронные сети.» Воронеж: ВГУ, 1994.
4. Уоссермен Ф. «Нейрокомпьютерная техника: теория и практика.» М.: Мир, 1992.
5. Мкртчян С.О. «Нейроны и нейронные сети. (Введение в теорию формальных нейронов)» М.: Энергия, 1971.
6. Гилл Ф., Мюррей У., Райт М. «Практическая оптимизация.» М.: Мир, 1985.
7. Вукалович М.П., Новиков И.И. «Термодинамика.» М., 1984
8. IAPWS «Industrial Formulation 1997 for the Thermodynamic Properties of Water and Steam. International Association for the Properties of Water and Steam» / Executive Secretary R.B. Dooley. Electric Power Research Institute. Palo Alto. CA 94304, USA
9. Катковский Е.А. и др., «О методах расчета на ЭВМ теплофизических свойств газов и жидкостей», Препринт ИАЭ, 3263/3, Москва, 1980 г.
10. Катковский Е.А. и др., «H₂O-пакет прикладных программ для расчета теплофизических свойств и их производных для воды и пара.» Препринт ИАЭ, 3344/16 Москва, 1980 г.
11. W. Wagner et al., "The IAPWS Industrial Formulation 1997 for the Thermodynamic Properties of Water and Steam", ASME J. Eng. Gas Turbines and Power, Vol. 122 (2000) pp. 150-182.

meilleurs paramètres de l'exactitude et de la vitesse de fonctionnement.

5. Les approches utilisées par les auteurs au cours de l'approximation des propriétés thermodynamiques, peuvent être généralisées également pour un autre support constant des calculs sur l'ordinateur des procès thermiques, hydrodynamiques et physico-chimiques. C'est actuel, par exemple, pour la création de la nouvelle génération des codes de calcul pour l'analyse des procès d'avarie et des ceux de transition chez IEN (les Installations de l'Énergie Nucléaire).

Le répertoire de la littérature

1. Kolmogorov A.N. «Sur la représentation des fonctions continues de quelques variables par les juxtapositions des fonctions continues du plus petit nombre des variables». L'exposé de l'Académie des sciences de l'URSS, 1956. Т. 108, №. 2 s 179-182.
2. Gorbagne A.N. «Le théorème généralisé approximant et les possibilités calculatoires des réseaux neuroniques». Le bureau de calcul du département de Sibérie de l'Académie des Sciences de la Russie à Krasnoïarsk, 1993.
3. Sourovtssev I.S., Klukine V.I., Pivovarova R.P. «Les réseaux neuroniques.» Voronej : L'Université d'état de Voronej, 1994.
4. F. Wosserman «Neurocomputer Technics.» La traduction russe : Moscou: maison d'éditions « Mir », 1992.
5. Mkrтчian S.O. «Les neurones et les réseaux neuroniques. (l'Introduction à la théorie des neurones formels)» Moscou, maison d'éditions: «Énerguiya », 1971.
6. Gill, P.E., Murray, W. and Wright, M.H. (1981) Practical Optimization, Academic Press.
7. Vukalovich M.P., Novikov I.I. «la Thermodynamique.» Moscou., maison d'éditions: «Énerguiya », 1984
8. IAPWS «Industrial Formulation 1997 for the Thermodynamic Properties of Water and Steam. International Association for the Properties of Water and Steam» / Executive Secretary R.B. Dooley. Electric Power Research Institute. Palo Alto. CA 94304, USA.
9. Katkovsky E.A. et autres auteurs, «Sur les méthodes du calcul sur l'ordinateur des propriétés thermophysiques des gaz et les liquides», la réimpression de l'Institut de l'Énergie atomique de I.V.Kurchatov, № 3263/3, Moscou, 1980.
10. Katkovsky E.A. et autres auteurs, «Le H₂O-paquet de programmes d'application pour le calcul des propriétés thermophysiques et de leurs dérivées pour l'eau et la vapeur», la réimpression de l'Institut de l'Énergie atomique de I.V.Kurchatov, № 3344/16 Moscou, 1980.
11. W. Wagner et al., "The IAPWS Industrial Formulation 1997 for the Thermodynamic Properties of Water and Steam", ASME J. Eng. Gas Turbines and Power, Vol. 122 (2000) pp. 150-182.

