

Методы и модели системного анализа

Обучение нейронной сети для прогнозирования свойств никелевых сплавов на основе генетического алгоритма*

О.С. Нургаянова¹, Н.И. Юсупова¹

¹ Федеральное государственное бюджетное образовательное учреждение высшего образования Уфимский государственный авиационный технический университет, г. Уфа, Россия

Аннотация. В статье рассматриваются вопросы прогнозирования свойств жаропрочных никелевых сплавов с монокристаллической структурой, применяемых в авиационной промышленности для изготовления лопаток газотурбинных двигателей. Приводятся постановка задачи, модели, методы и алгоритмы решения с помощью технологии искусственных нейронных сетей. Даны подробные описания алгоритмов для решения задачи прогнозирования свойств сплавов с помощью нейронных сетей – алгоритм обратного распространения ошибки и генетический алгоритм обучения; сравнение результатов прогнозирования свойств сплавов с помощью нейронных сетей, обученных методом обратного распространения ошибки и генетическим алгоритмом; особенности программной реализации и обсуждение полученных результатов.

Ключевые слова: нейронные сети, жаропрочные никелевые сплавы, обратное распространение ошибки, генетический алгоритм обучения.

DOI: 10.14357/20790279190403

Введение

Известно, что процесс разработки новых материалов достаточно трудоемкий и включает в себя множество последующих испытаний на исследование различных характеристик. Использование современных методов математического моделирования и прогнозирования позволит ускорить процесс разработки моделей «состав-свойство» и предусмотреть возможные диапазоны попадания исследуемых характеристик.

В задаче прогнозирования свойств никелевых сплавов на основе использования технологии искусственных нейронных сетей важным является выбор метода обучения сети. Основным алгоритмом обучения искусственных нейронных сетей – метод обратного распространения ошибки – известен

уже достаточно давно и хорошо исследован, но, тем не менее, и сегодня обучение нейронных сетей по-прежнему является нетривиальной и в значительной степени творческой задачей [1,2]. При решении реальных задач создаются нейронные сети со сложной структурой, в связи с чем, их обучение требует гораздо большего времени [4, 8].

Ранее [7], задача, связанная с прогнозированием свойств сплавов, решалась с помощью многослойного перцептрона с количеством входных нейронов равных количеству химических элементов, с обучением на основе метода обратного распространения ошибки.

В данной статье рассматривается генетический алгоритм обучения нейронной сети для решения задачи прогнозирования комплексного свойства – жаропрочности никелевых сплавов и требования к программному модулю, реализующему

* Исследование частично поддержано грантом РФФИ 18-07-00193-а.

цему выбранную архитектуру сети и метод обучения на основе генетического алгоритма. Выбор данного метода обусловлен возможностью поиска наилучшего значения жаропрочности среди большого числа комбинаций химических элементов при наличии внутренних латентных зависимостей в переменных.

1. Постановка задачи

Дано множество объектов – сплавов, принадлежащих одному классу, которое описывается набором векторов $\langle x_1, x_2, \dots, x_n \rangle = y$, где $\langle x_1, x_2, \dots, x_n \rangle$ – вектор признаков – состав сплава, представленный процентным содержанием химических элементов, y – комплексная характеристика объекта – жаропрочность, зависящая от вектора $\langle x_1, x_2, \dots, x_n \rangle$ – состава сплава. Вид зависимости между входными значениями и выходными не известен.

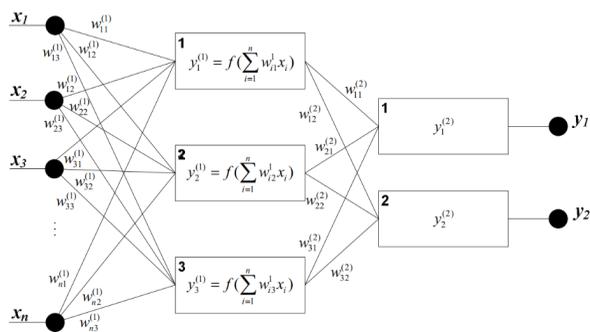


Рис. 1. Двуслойная нейронная сеть с весовыми коэффициентами факторов

Таким образом, имеем оптимизационную задачу, в которой требуется подобрать такой вектор признаков, которых бы максимизировал функцию жаропрочности – y . На рисунке 1 показана выбранная для решения задачи прогнозирования архитектура нейронной сети.

2. Анализ известных работ

Разработке и обучению нейронной сети методом обратного распространения ошибки для анализа и интерпретации данных посвящен ряд работ [2, 7, 8, 12, 13, 14, 15], в которых так же говорится, что вычислительные эксперименты выявили следующие недостатки: возможность преждевременной остановки из-за попадания в область локального минимума; необходимость многократного предъявления всего обучающего множества для получения заданного качества распознавания; отсутствие приемлемых оценок времени обучения. Исследо-

вания показали, что исключить данные недостатки и повысить скорость обработки данных позволяет генетический алгоритм, используемый на этапе обучения нейронной сети. Разработке методики его применения для прогнозирования свойств никелевых сплавов и посвящена данная статья.

3. Интерпретация терминов генетического алгоритма для задачи обучения нейронной сети

Генетический алгоритм является самым известным на данный момент представителем эволюционных алгоритмов и по своей сути является алгоритмом для нахождения глобального экстремума многоэкстремальной функции. Суть его заключается в параллельной обработке множества альтернативных решений, при этом поиск концентрируется на наиболее перспективных из них [1, 2].

При описании генетических алгоритмов используются определения, заимствованные из генетики. Говоря о популяции особей, в качестве базовых понятий применяются термины: ген, хромосома, генотип, фенотип, аллель.

В терминах обучения нейронной сети основные понятия генетического алгоритма интерпретируются следующим образом:

- ген – весовой коэффициент нейронной сети;
- хромосома – набор генов (т.е. весовых коэффициентов нейронной сети, считываемых в определенном порядке сверху вниз, справа налево). Каждая хромосома является возможным решением (т.е. таким набором весовых коэффициентов, которые лучше подходят для установления зависимости «состав-свойство»);
- популяция – множество хромосом, вариантов наборов весовых коэффициентов;
- эпоха – итерация, соответствующая созданию нового поколения хромосом.

Хромосомы являются основными сущностями, над которыми в определенном порядке в пределах одной эпохи проводятся следующие операции:

- скрещивание – создание с определенной степенью вероятности новой хромосомы из генов двух других и добавление ее в популяцию;
- мутация – изменение с определенной степенью вероятности значения произвольного гена любой хромосомы и добавление ее в популяцию;
- приспособление – удаление из популяции хромосом (т.е. наборов весовых коэффициентов), показавших худший результат при распознавании.

На рисунке 1 показана выбранная архитектура нейронная сеть, весовыми коэффициентами

которой (производя обход сверху вниз, справа на- лево) заполняются хромосомы.

4. Разработка и реализация алгоритма

Перед тем как приступить к обучению ней- ронной сети генетическим алгоритмом, необходи- мо сформировать начальную популяцию, то есть какое-то количество матриц весов $W^{нач}$. Под по- пуляцией будем понимать набор хромосом (реше- ний), которые формируются случайным образом, но с учетом условия $-1 \leq w_{ij} \leq 1$. Данная матрица имеет размер $[N \times M]$. В процессе обучения нейроны могут быть удалены из сети вследствие их малого влияния на работу всей сети. Следовательно, раз- мер матрицы весов может уменьшаться. В случае удаления j -ого нейрона из сети, в матрице весов вычеркивается j -ый столбец и j -ая строка. Далее работа ведется уже с новой матрицей. Рассмотрим основные операторы генетического алгоритма.

Оператор скрещивания, иногда называемый кроссинговером, является основным генетическим оператором (Рисунок 2), за счет которого произво- дится обмен частями хромосом между двумя (мо- жет быть и больше) хромосомами в популяции. Оператор скрещивания может быть одноточечным или многоточечным. Одноточечным называется скрещивание, если родительские хромосомы раз- резаются только в одной случайной точке, многото- чечным – в случае двух и более точек разреза. Реа- лизация многоточечного скрещивания может быть происходит двумя способами: либо точек разрыва меньше, чем генов в хромосоме, либо – если дли- на хромосомы L бит, то число точек разрыва равно $(L-1)$ и потомки наследуют биты следующим об- разом: первому потомку достаются нечетные биты

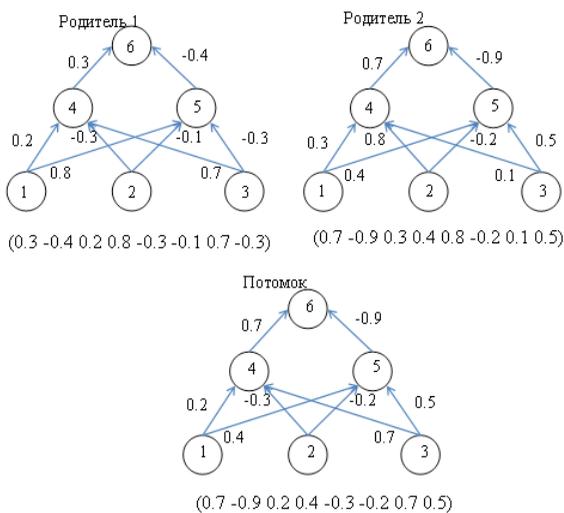


Рис. 2. Работа оператора скрещивания

первого родителя и четные биты второго; у второго же потомка все наоборот.

Кроме описанных способов скрещивания су- ществует ещё однородное скрещивание. Его осо- бенность заключается в том, что значение каждого бита в хромосоме потомка определяется случай- ным образом из соответствующих битов родите- лей. Для этого вводится некоторая величина $-1 \leq P_o \leq 1$, и если случайное число больше P_o , то на n -ю позицию первого потомка попадает n -й бит перво- го родителя, а на n -ю позицию второго – n -й бит второго родителя. В противном случае к первому потомку попадает бит второго родителя, а ко вто- рому – первого. Такая операция проводится для всех битов хромосомы. Вероятность скрещивания самая высокая среди генетических операторов и равна обычно 60% и более.

Работа оператора мутации представляет со- бой стохастическое изменение части хромосом. Данная процедура необходима для «выбивания» популяции из локального экстремума и способ- ствует защите алгоритма от преждевременной схо- димости. Достигается это за счет того, что каждый ген строки, которая подвергается мутации, с малой вероятностью меняется на другой ген (к весу до- бавляется случайная величина между -1.0 и 1.0).

Так же как и скрещивание, мутация прово- дится не только по одной случайной точке. Можно выбирать некоторое количество точек в хромосоме для изменения, причем их число также может быть случайным.

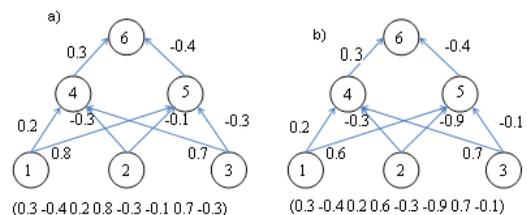


Рис. 3. Работа оператора мутации

Также можно изменять сразу некоторую груп- пу подряд идущих точек. Вероятность мутации значительно меньше вероятности скрещивания и редко превышает 1%. Выше на рисунке 3 показана работа оператора мутации – а) состояние объекта до начала мутации; б) – состояние после мутации.

Селекция хромосом заключается в выбо- ре (по рассчитанным на втором этапе значениям функции приспособленности) тех хромосом, ко- торые будут участвовать в создании потомков для следующей популяции, т.е. для очередного поколе- ния. Такой выбор производится согласно принци- пу естественного отбора, по которому наибольшие

шансы на участие в создании новых особей имеют хромосомы с наибольшими значениями функции приспособленности. Существуют различные методы селекции. Наиболее популярным считается так называемый метод рулетки, который свое название получил по аналогии с известной азартной игрой. Каждой хромосоме может быть сопоставлен сектор колеса рулетки, величина которого устанавливается пропорциональной значению функции приспособленности данной хромосомы. Поэтому чем больше значение функции приспособленности, тем больше сектор на колесе рулетки. Все колесо рулетки соответствует сумме значений функции приспособленности всех хромосом рассматриваемой популяции.

Каждой хромосоме, обозначаемой Ch_i для $i=1, 2, \dots, N$ (где N – численность популяции) соответствует сектор колеса $v(Ch_i)$, выраженный в процентах согласно формуле:

$$v(Ch_i) = P_s(Ch_i) / 100\%$$

$$P_s(Ch_i) = F(Ch_i) / \sum_{i=1}^N F(Ch_i),$$

причем $F(Ch_i)$ – значение функции приспособленности хромосомы, а $P_s(Ch_i)$ – вероятность селекции хромосомы (Ch_i). Селекция хромосомы может быть представлена как результат поворота колеса рулетки, поскольку «выигравшая» (т.е. выбранная) хромосома относится к выпавшему сектору этого колеса. Очевидно, что чем больше сектор, тем больше вероятность «победы» соответствующей хромосомы. Поэтому вероятность выбора данной хромосомы оказывается пропорциональной значению ее функции приспособленности. Если всю окружность колеса рулетки представить в виде цифрового интервала $[0, 100]$, то выбор хромосомы можно отождествить с выбором числа из интервала $[A, B]$, где A и B обозначают соответственно начало и окончание фрагмента окружности, соответствующего этому сектору колеса; очевидно, что $1 \leq A < B \leq 100$. В этом случае выбор с помощью колеса рулетки сводится к выбору числа из интервала $[0, 100]$, которое соответствует конкретной точке на окружности колеса. Существуют также другие методы селекции. В результате процесса селекции создается родительская популяция, также называемая родительским пулом с численностью N , равной численности текущей популяции.

В процессе работы алгоритма все указанные выше операторы применяются многократно и ведут к постепенному изменению исходной популяции. Поскольку операторы отбора, скрещивания, мутации и редукции по своей сути направлены на улучшение каждой отдельной особи, то результатом их

работы является постепенное улучшение популяции в целом. В этом и заключается основной смысл работы генетического алгоритма – улучшить популяцию решений по сравнению с исходной.

После завершения работы генетического алгоритма из конечной популяции выбирается та особь (сплав), которая дает экстремальное (максимальное или минимальное) значение целевой функции и является, таким образом, результатом работы генетического алгоритма. За счет того, что конечная популяция лучше исходной, полученный результат представляет собой улучшенное решение.

Порядок обучения нейронной сети с применением генетического алгоритма представлен на рисунке 4.

Критерием останова работы алгоритма является достижение целевой функции оптимизации:

$$W = \sum \Delta_j \rightarrow \max, j \in [0, \dots, K].$$

Рассмотрим квадратичную сумму для одного примера из выборки:

$$\Delta_j = \sum (Y_i^{id} - Y_i^{real})^2, j \in [1, \dots, m],$$

где Y_i^{id} – значение i -ого выхода нейронной сети на обучающем примере данной выборки, Y_i^{real} – значение i -ого выхода, полученное в результате реальной работы нейронной сети. На следующем шаге рассчитывается квадратичная сумма ошибки для каждого из K примеров выборки.

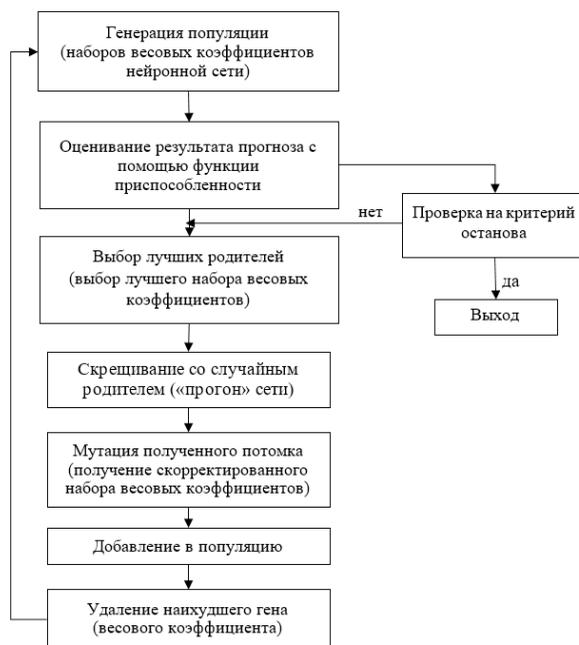


Рис. 4. Схема обучения нейронных сетей с помощью генетического алгоритма

5. Анализ результатов

Сравнивалась работа нейронной сети, обученной с помощью генетического алгоритма с сетью той же архитектуры, обученной алгоритмом обратного распространения ошибки. На вход двуслойной нейронной сети, состоящей из 256 входных элементов (по 3 нейрона в первом и во втором слое и 1 нейрон в выходном) подавался обучающий набор векторов; обучение считалось законченным, когда во всем обучающем наборе результирующий выходной сигнал совпадал с требуемым эталонным выходным сигналом.

Вычислительный эксперимент показал:

- Для обучения нейронной сети с помощью генетического алгоритма достаточно выполнить 100 генераций, в то время при обучении алгоритмом обратного распространения требуется более 500 итераций (где одна итерация – это полный пересчет всех обучающих данных: весовых коэффициентов, погрешности и выходных значений).
- Две генерации в генетическом алгоритме эквивалентны одной итерации алгоритма обратного распространения, так как обратное распространение состоит из двух частей прямого прохода (вычисления выхода сети и ошибки) и обратного прохода (изменения весов). Генетические алгоритмы выполняют только первую часть. Таким образом, две генерации занимают меньше времени, чем вычисления единственной итерации алгоритма обратного распространения.

В целом, генетические алгоритмы значительно выигрывают по сравнению со схемой обратного распространения, позволяя получать лучшие векторы весовых коэффициентов за существенно меньшее время.

Заключение

Дальнейшие направления исследований связаны с разработкой более эффективных алгоритмов обучения нейронных сетей, введения в архитектуру сети нейронов по «переменным протекции» и нейрона смещения [2, 7, 8]. Отдельным направлением исследований можно выделить изучение фазового состава сплавов с помощью методов Data Mining и методов хемоинформатики.

Литература

1. Рутковская Д., Пилинский М., Рутковский Л. «Нейронные сети, генетические алгоритмы и нечеткие системы», М.: Горячая линия - Телеком, 2006.
2. Хайкин С. «Нейронные сети: полный курс», Пер. с англ.- М.: Издательский дом «Вильямс», 2006.
3. Дюк В., Самойленко А. «Data Mining: учебный курс (+CD)», СПб: Питер, 2001.
4. Круглов В., Борисов В. «Искусственные нейронные сети. Теория и практика», М.: Горячая линия–Телеком, 2001.
5. Superalloys II: High-Temperature Materials for Aerospace and Industrial Power 2nd Edition Chester T. Sims, Norman S. Stoloff, William C. Hagel © 1987 by John Wiley & Sons Inc.
6. Рахманкулов М.М., Паращенко В.М. Технология литья жаропрочных сплавов. М: Интермет-Инжиниринг, 2000 – 464 с.
7. Нургаянова О.С. Автоматизированное проектирование литейных жаропрочных никелевых сплавов на основе методов искусственного интеллекта: дис. канд. техн. наук. — Уфа, 2006.
8. Нургаянова О.С. Применение искусственных нейронных сетей в задачах классификации многокомпонентных сплавов // Proceedings of the 6th All-Russian Scientific Conference «Information Technologies for Intelligent Decision Making Support», Volume 3, May 28-31, Ufa- Stavropol, Russia, 2018
9. Айвазян С.А. Анализ данных, прикладная статистика и построение общей теории автоматической классификации// Методы анализа данных/ Пер. с фр. — М.: Финансы и статистика, 1985.
10. Mohammed J. Zaki, Wagner Meira Jr. Data Mining and Analysis. Fundamental Concepts and Algorithms. – Cambridge University Press, 2014. 593 с.
11. Rashid T. Make Your Own Neural Network. CreateSpace, 2016. — 222 p.
12. Groumos, P.P. (2016). Deep Learning vs. Wise Learning: A Critical and Challenging Overview. In IFAC Intern. Conf. on International Stability, Technology and Culture, TECIS 2016, Durres, Albania.
13. Юсупова Н.И., Ризванов Д.А., Сметанина О.Н., Еникеева К.Р. Модели представления знаний для поддержки принятия решений при управлении сложными системами в условиях неопределенности и ресурсных ограничений / Information technologies for intelligent decision making support (ITIDS'2016) Proceedings of the 4th International Conference. 2016
14. Обработка слабоструктурированной информации на основе методов искусственного интеллекта: монография / Н. И. Юсупова, Д.Р. Богданова, М. В. Бойко.– М.: Инновационное машиностроение, 2016. – 250 с.

15. *Гузаиров М.Б., Юсупова Н.И., Сметанина О.Н. и др.* Методические аспекты искусственного интеллекта (монография) / Уфимский государственный авиационный технический университет, Научный совет РАН по методологии искусственного интеллекта, Рецензенты: профессор В.М. Картак, профессор Д.Е. Сорокин. Москва, 2014.

Нургаянова Ольга Сергеевна. Федеральное государственное образовательное бюджетное учреждение высшего образования Уфимский государственный авиационный технический университет (ФГБОУ ВО УГАТУ) г. Уфа, Россия. Кандидат технических наук. Доцент кафедры Вычислительной математики и кибернетики ФГБОУ ВО УГАТУ. Количество печатных работ: 20. Область научных интересов: методы искусственного интеллекта, алгоритмы обучения нейронных сетей, информационные технологии. E-mail: onurgayanova@yandex.ru

Юсупова Нафиса Исламовна. Федеральное государственное образовательное бюджетное учреждение высшего образования Уфимский государственный авиационный технический университет (ФГБОУ ВО УГАТУ) г. Уфа, Россия. Доктор технических наук, профессор. Заведующий кафедрой Вычислительной математики и кибернетики ФГБОУ ВО УГАТУ. Количество печатных работ: более 200. Область научных интересов: системный анализ, методы искусственного интеллекта и принятия решений. E-mail: yussupova@ugatu.ac.ru

Training a neural network to predict the properties of nickel alloys with a genetic algorithm

O.S. Nurgayanova¹, N.I. Yussupova¹

¹Ufa State Aviation Technical University, Ufa, Russia

Abstract. The article presents questions of predicting the properties of heat-resistant nickel alloys with a single crystal structure used in the aviation industry for the manufacture of gas turbine engine blades. The article presents the statement of the problem, models, methods and algorithms for solving with technology of artificial neural networks. Detailed descriptions of the algorithms for solving the problem of predicting the properties of alloys using neural networks are given – back propagation and a genetic algorithms; comparing the results of predicting the properties of alloys with neural networks, trained by the method of back propagation and the genetic algorithms; features of software implementation and discussion of the results.

Keywords: *neural networks, heat-resistant nickel alloys, back propagation algorithm, genetic algorithm for learning neural networks.*

DOI: 10.14357/20790279190403

References

1. *Rutkovskaya D., Pilinsky M., Rutkovsky L.* Neural networks, genetic algorithms and fuzzy systems, M.: Goryachaya liniya – Telecom, 2006.
2. *Haykin Simon.* 2006. Neural networks and learning machines / Simon Haykin.—2rd ed. p. cm
3. *Dyuk V., Samoilenko A.* Data Mining: training course (+ CD)”, St. Petersburg: Peter, 2001.
4. *Kruglov V., Borisov V.* Iskusstvennyye neyronnyye seti. Teoriya i praktika [Artificial neural networks. Theory and Practice], Moscow: Goryachaya liniya – Telecom, 2001.
5. *Superalloys II: High-Temperature Materials for Aerospace and Industrial Power* 2nd Edition Chester T. Sims, Norman S. Stoloff, William C. Hagel © 1987 by John Willey & Sons Inc.
6. *Rakhmankulov M.M., Parashchenko V.M.* Tekhnologiya lit'ya zharoprochnykh spлавov [Technology for casting heat resistant alloys]. M: Internet Engineering, 2000 - 464 p.
7. *Nurgayanova O.S.* 2006 Avtomatizirovannoye proyektirovaniye liteynykh zharoprochnykh nikelovykh spлавov na osnove metodov iskusstvennogo intellekta [Computer-aided design of cast heat-resistant nickel alloys using artificial intelligence methods]. PhD Thesis. Ufa: USATU 16 p.
8. *Nurgayanova O.S.* Primeneniye iskusstvennykh neyronnykh setey v zadachakh klassifikatsii

- mnogokomponentnykh splavov [The use of artificial neural networks for the classification of multicomponent alloys] // Proceedings of the 6th All-Russian Scientific Conference «Information Technologies for Intelligent Decision Making Support», Volume 3, May 28-31, Ufa- Stavropol, Russia, 2018.
9. *Ayvazyan S.A.* Data analysis, statistics and the construction of a general theory of automatic classification // *Methods of data analysis / Per. from French - M.: Finance and Statistics, 1985.*
 10. *Mohammed J. Zaki, Wagner Meira Jr.* (2014) *Data Mining and Analysis. Fundamental Concepts and Algorithms.* – Cambridge University Press, 593 p.
 11. *Rashid T.* *Make Your Own Neural Network. CreateSpace, 2016. — 222 p.*
 12. *Groumpos, P. P.* (2016). *Deep Learning vs. Wise Learning: A Critical and Challenging Overview.* In *IFAC Intern. Conf. on International Stability, Technology and Culture, TECIS 2016, Durres, Albania.*
 13. *Yussupova N.I., Rizvanov D.A., Smetanina O.N., Enikeeva K.R.* (2016) *Knowledge modeling to support decision making in managing complex conditions in the face of uncertainty and resource constraints / Information technologies for intelligent decision making support (ITIDS'2016) Proceedings of the 4th International Conference.*
 14. *Yussupova N., Bogdanova D., Boyko M.* (2016) *The processing of poorly structured information based on artificial intelligence methods: monograph / M.: Innovative Engineering, 250 p.*
 15. *Guzairov M.B., Yussupova N.I., Smetanina O.N. et al.* (2014) *Methodological aspects of artificial intelligence (monograph) / Ufa State Aviation Technical University, Scientific Council of the Russian Academy of Sciences on the methodology of artificial intelligence, Reviewers: Professor Kartak V.M., professor Sorokin D.E. Moscow.*

Nurgayanova O.S. Candidate of Technical Sciences, Associate Professor, Department of Computational Mathematics and Cybernetics. Ufa State Aviation Technical University, 12 K. Marx str., Ufa, 450000, Russia. E-mail: onurgayanova@yandex.ru

Yussupova N.I. Professor, Chief of Department of Computer Science and Robotics. Ufa State Aviation Technical University, 12 K. Marx str., Ufa, 450000, Russia. E-mail: yussupova@ugatu.ac.ru