

ИНФОРМАТИКА,
ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА
И УПРАВЛЕНИЕ

INFORMATICS,
COMPPUTER ENGINEERING
AND CONTROL

УДК 004.852

DOI: 10.17212/1814-1196-2020-1-107-118

Классификация станков, описываемых большим количеством признаков, на основе нейронной сети распознавания и редукции пространства признаков^{*}

С.С. СОСИНСКАЯ^а, С.А. РОГАЧЕВА^б

*664074, РФ, г. Иркутск, ул. Лермонтова, 83, Иркутский национальный исследова-
тельский технический университет*

^а sosinskaya@mail.ru ^б pyatsot@bk.ru

В последние годы растет актуальность классификации информации с использованием нейронных сетей. Это связано с тем, что объемы данных, которые необходимо обрабатывать, становятся с каждым днем всё больше. Практически все современные программные комплексы характеризуются большим разнообразием взаимодействующих программных модулей, что, в свою очередь, увеличивает сложность обработки данных. На этом этапе разработчики программных продуктов используют различные методы.

Нейронные сети позволяют достигать максимально высоких скоростей и высокой точности в работе с большим объемом информации по сравнению с другими методами классификации и обработки данных. В связи с этим задача разработки методов классификации становится всё более актуальной.

В статье рассматривается подход к классификации выборки моделей станков, описываемых большим количеством признаков. В качестве инструмента для проведения классификации используется два наиболее известных типа нейронных сетей: многослойный перцептрон и сеть распознавания. В связи с тем, что ни на одном типе сети на полном наборе признаков не было получено высокое качество классификации, авторы применили метод главных компонент (РСА) для редукции пространства признаков, что повлекло за собой существенное повышение качества классификации.

Разработанный подход может применяться для классификации моделей станков, не представленных в выборке. Кроме того, статья иллюстрирует тот факт, что выбор метода классификации в значительной степени зависит от вида предметной области и характера выборки.

Программная реализация основана на использовании системы MATLAB, которая предоставляет множество инструментов и методов для подготовки, анализа и визуализации данных.

Ключевые слова: классификация, нейронные сети, разработка методов, многослойный перцептрон, сеть распознавания, метод главных компонент, метод РСА, редукция пространства признаков

^{*} *Статья получена 30 октября 2019 г.*

ВВЕДЕНИЕ

Интеллектуальный анализ данных является одним из самых актуальных направлений исследований в современном мире, так как он позволяет упростить и рационализировать работу со значительным объемом данных различных типов (количественные, категориальные, текстовые).

Данные – это конкретная форма представления содержания информации. Например, информацию о результатах наблюдения за температурой окружающей среды можно представить в виде числового массива (таблицы), но можно и в виде графика, и в виде текстового описания посредством некоторого языка [14].

Классификация данных – определение правил и алгоритмов для группирования множества объектов на основе выборки данных. Классифицировать объект означает указать номер (или наименование) класса, к которому он относится. Компьютерная классификация позволяет отыскивать закономерности в данных, которые человек не находит [8].

Каждая группа (класс, кластер) характеризуется определенными признаками, которыми наделены все содержащиеся в нем объекты. Число классов заранее известно из каких-либо понятийных для предметной области соображений.

Анализ содержимого баз данных большого объема является очень трудоемким. Исходя из этого целесообразно реализовать ряд методов, позволяющих автоматизировать процесс классификации данных [6, 13].

Для решения этой задачи требуется выбрать наиболее эффективный метод классификации. Для проведения интеллектуального анализа данных в статье использовалась база данных станков [1]. Эта база была выбрана из-за ее достаточно большого объема, а именно, она содержит 691 модель станков, каждая из которых характеризуется 185 признаками.

Полученные результаты позволят классифицировать модели станков, не представленные в выборке, а также проиллюстрировать тот факт, что выбор метода классификации в значительной степени зависит от вида предметной области и характера выборки.

1. ПОСТАНОВКА ЗАДАЧИ

Каждая модель станка в базе данных относится к определенной группе. Всего имеется 9 групп:

- токарные;
- сверлильные и расточные;
- шлифовальные, полировальные, доводочные;
- комбинированные;
- зубо- и резьбообрабатывающие;
- фрезерные;
- строгальные, долбежные, протяжные;
- разрезные;
- обрабатывающие центры.

Общие признаки, присущие моделям всех групп, следующие:

- наибольший диаметр обрабатываемого изделия, мм;
- минимальная частота вращения;
- максимальная частота вращения;
- минимальная мощность двигателя, кВт;
- максимальная мощность двигателя, кВт;
- длина обрабатываемого изделия, мм;
- длина станка, мм;
- ширина станка, мм;
- высота станка, мм;
- вес станка, кг.

Кроме того, каждая группа станков имеет определенный набор признаков. Например, станки токарной группы имеют следующие признаки:

- максимальное количество инструментов;
- диаметр планшайбы / шпинделя, мм;
- масса установленного изделия;
- подача, об/мин;
- перемещение суппорта продольное, мм/об;
- перемещение суппорта поперечное, мм/об;
- диаметр стола, мм.

Для группы «Сверлильные и расточные станки» заданы признаки:

- максимальный ход шпинделя, мм;
- максимальное расстояние от шпинделя до колонны станка, мм;
- минимальный и максимальный диапазон резьбы, мм;
- максимальный диаметр растачивания, мм;
- максимальный крутящий момент на шпинделе, мм.

Таким образом, создана единая таблица всех признаков (в количестве 185; вместо отсутствующих признаков некоторых моделей указан 0).

Для исключения данных из разных шкал следует нормализовать (привести к виду $[0,1]$) признаки таблицы, в которой будут исходные данные для классификации и анализа, необходимого для решения задачи.

2. ЭТАПЫ РЕШЕНИЯ ЗАДАЧИ КЛАССИФИКАЦИИ НА ОСНОВЕ НЕЙРОННЫХ СЕТЕЙ

В целом проблема классификации состоит из двух частей: обучения и распознавания. Обучение осуществляется путем показа отдельных объектов с указанием их принадлежности определенному классу. Это обучение с учителем. В результате обучения нейронная сеть должна приобрести способность реагировать одинаково на все объекты одного класса и различным образом – на все объекты различных классов. Очень важно, что процесс обучения должен происходить только путем показов конечного числа объектов без каких-либо других подсказок [4]. После завершения обучения можно распознавать (то есть классифицировать) другие объекты [15].

Для получения лучшего качества классификации исследовались два типа нейронных сетей: многослойный перцептрон и сеть распознавания.

Многослойный перцептрон – это класс искусственных нейронных сетей прямого распространения, состоящих из нескольких слоев:

- множества входных узлов, которые образуют входной слой;
- нескольких скрытых слоев нейронов;
- выходного слоя нейронов.

В каждом слое имеются веса и смещение, которые поступают на вход сумматора; выход сумматора поступает на вход функции активации, выход которой – либо вход следующего слоя, либо выход сети [2, 12].

В скрытом слое у многослойного перцептрона используется функция активации – гиперболический тангенс (Tansig), вычисляемый по формуле

$$f(x) = \tanh(x) = \frac{2}{1 + e^{-2x}} - 1. \quad (1)$$

В выходном слое многослойного перцептрона используется линейная функция (*linear*) активации (рис. 1). Это позволяет получать на выходе одно из значений, понимаемых как номер класса.

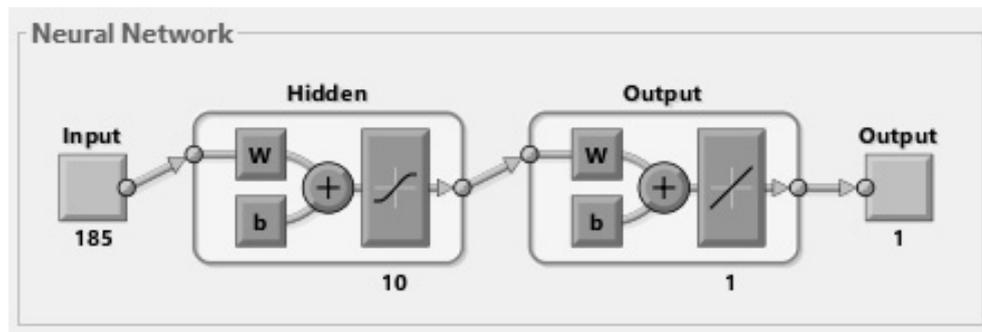


Рис. 1. Многослойный перцептрон с линейной функцией активации

Fig. 1. A multilayer perceptron with the linear activation function

Сеть распознавания (рис. 2) – это частный случай сети прямого распространения, которая может быть обучена классификации входных данных в соответствии с целевыми классами. Эти классы для сетей распознавания представлены векторами, содержащими нулевые значения, за исключением единичного значения в элементе i , где i – номер класса, к которому относится наблюдение [3].

В выходном слое сети распознавания используется функция активации softmax:

$$a_i = \frac{e^{Z_i}}{\sum_{j=1}^n e^{Z_j}}, \quad (2)$$

где Z_i – значение на выходе i -го нейрона; n – общее число нейронов в слое. Функция преобразует вектор Z размерности n в вектор a той же размерности, где каждая координата a_i полученного вектора представлена вещественным числом, находящимся в интервале $[0 \dots 1]$, и сумма координат равна 1 [11].

В выходном слое находится количество нейронов, соответствующее количеству классов (в нашем случае 9 групп станков).

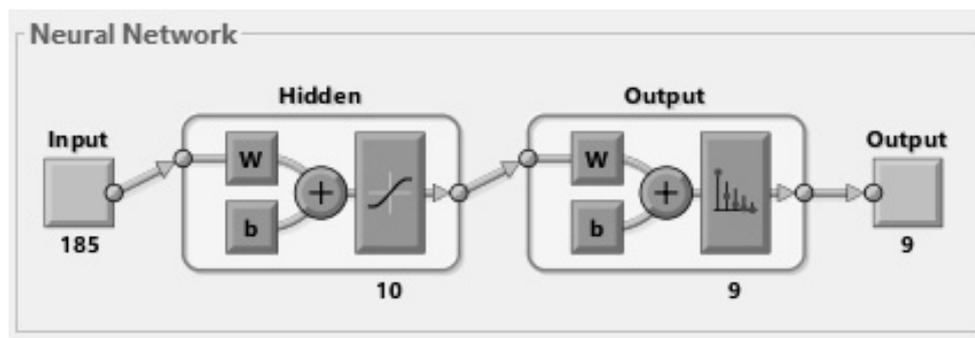


Рис. 2. Сеть распознавания

Fig. 2. A recognition network

При использовании нейросети исходные данные обычно разбиваются на обучающую и тестирующую выборки. *Обучающая выборка* используется на этапе обучения той или иной модели, тогда как *тестирующая выборка* служит для получения оценки прогнозных свойств модели на новых данных, т. е. данных, которые не были использованы для обучения модели. Для рассматриваемого набора моделей станков обучающая выборка составляла 75 % от объема исходных данных.

Число итераций обучения также подбиралось на основе графика проверки работоспособности [10]. Он имеет вид, приведенный на рис. 3, и показывает зависимость среднеквадратичной ошибки от номера итерации. Как видно, показатель ошибки в процессе обучения сокращается, но может увеличиваться на множестве проверочных данных. На графике выделена точка, соответствующая рекомендуемому количеству итераций (рис. 3); при дальнейшем увеличении числа итераций происходит так называемое переобучение (явление, когда построенная модель хорошо объясняет примеры из обучающей выборки, но относительно плохо работает на примерах, не участвовавших в обучении) [5, 8].

Другим способом избавиться от переобучения является сокращение размерности пространства признаков. В статье для этой цели был использован метод главных компонент (РСА), который заключается в сокращении размерности пространства признаков с минимальной потерей полезной информации [7, 9].

Метод главных компонент применяется к данным, записанным в виде матрицы X – прямоугольной таблицы чисел размерностью n строк и m столбцов.

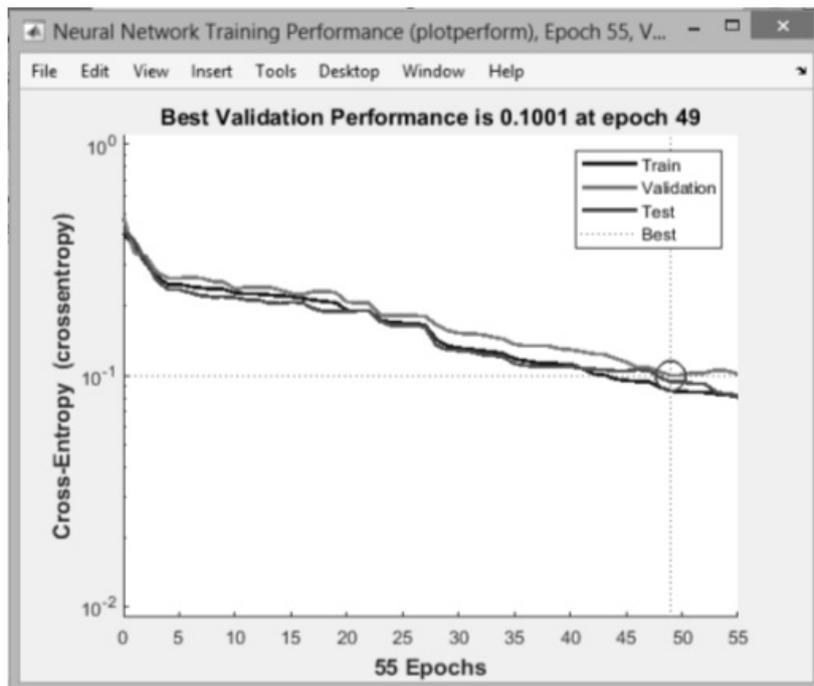


Рис. 3. График процесса обучения нейронной сети

Fig. 3. Graph of the neural network learning process

В методе главных компонент используются новые переменные T_a ($a = 1, \dots, A$), являющиеся линейной комбинацией исходных переменных x_j ($j = 1, \dots, m$). С помощью этих новых переменных матрица X представляется в виде произведения двух матриц T и P :

$$X = TP^T + E = \sum_{a=1}^A t_a P_a^t + E. \quad (3)$$

Матрица T называется матрицей счетов (scores), ее размерность $(i \times a)$.

Матрица P называется матрицей нагрузок (loadings), ее размерность $(a \times j)$.

E – матрица остатков размерностью $(i \times j)$.

Новые переменные t_a называются главными компонентами. Число столбцов в матрице T , и p_a в матрице P равно A (число главных компонент).

Важным свойством PCA является ортогональность (независимость) главных компонент.

Поэтому матрица T не перестраивается при увеличении числа компонент, а к ней просто добавляется еще один столбец, соответствующий новой главной компоненте. Это же происходит с матрицей нагрузок P .

Результатом PCA-моделирования являются величины \hat{x}_c – оценки, найденные по модели, построенной на обучающем наборе x_c . Результатом

проверки служат величины \hat{x}_t – оценки проверочных значений x_t , вычисленные по той же модели, но как новые образцы. Отклонение оценки от проверочного значения вычисляют как матрицу остатков:

- по обучающей выборке, формула (4):

$$E_c = x_c - \hat{x}_c \quad (4)$$

- по тестирующей выборке, формула (5):

$$E_t = x_t - \hat{x}_t. \quad (5)$$

При выполнении метода главных компонент вычисляются:

- матрица главных компонент (нагрузки);
- счета главных компонент – представление исходной матрицы наблюдений в координатах главных компонент;
- процент от общей дисперсии, объясняемой каждой главной компонентой.

Построив график зависимости суммарного процента от координат (рис. 4), можно увидеть, что для объяснения 90 % дисперсии нужно примерно 40 первых главных компонент.

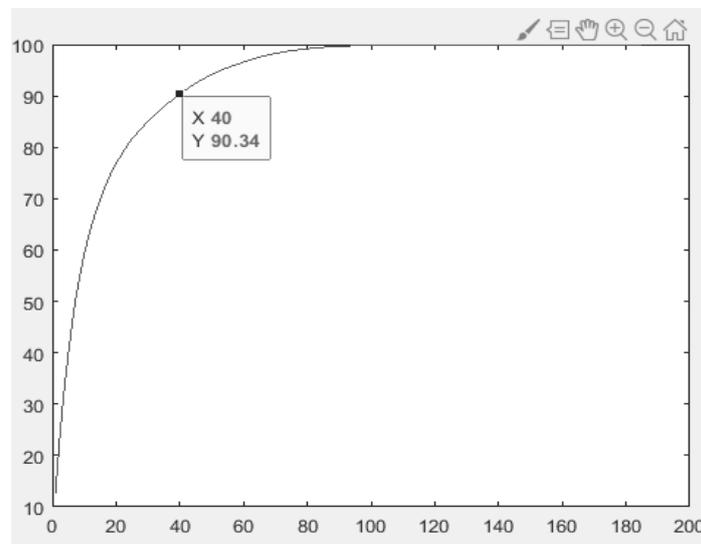


Рис. 4. График зависимости суммарного процента дисперсии от номеров координат

Fig. 4. Graph of the dependence of the total variance percentage on the coordinate numbers

Обучение и классификация по обученным сетям повторялись в пространстве размерности 40.

На рис. 5 показана структура многослойного персептрона для сокращенного пространства признаков.

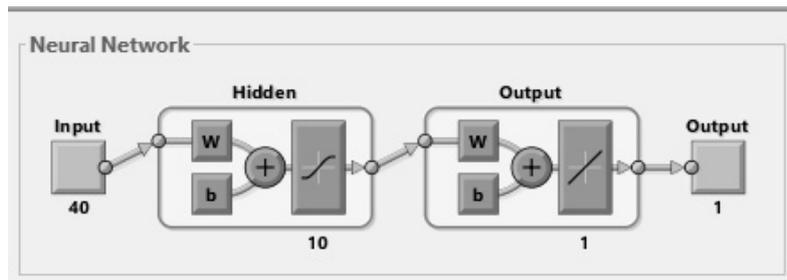


Рис. 5. Структура многослойного персептрона для сокращенного пространства признаков

Fig. 5. Structure of a multilayer perceptron for a reduced feature space

Структура сети распознавания для сокращенного пространства признаков представлена на рис. 6.

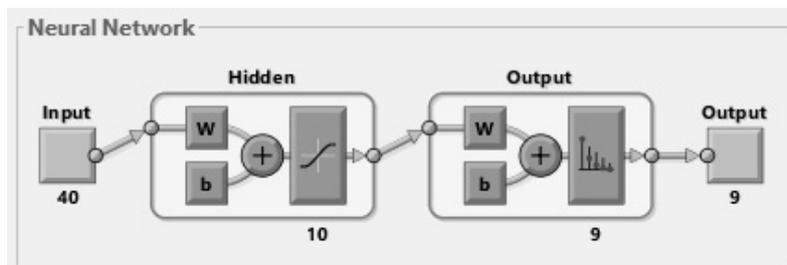


Рис. 6. Структура сети распознавания для сокращенного пространства признаков

Fig. 6. Structure of the recognition network for a reduced feature space

Результаты классификации по сетям с сокращенным пространством признаков оценивались графически путем отображения матриц ошибок соответственно по обучающей и тестирующей выборкам для многослойного персептрона и сети распознавания (рис. 7а–7г).

True class \ Predicted class	0	1	2	3	4	5	6	7	8	9
0	4									
1	31									
2	6	39			2					
3	2	1	118	2						
4			4	9	4		1	1		
5				1	87	1	1			
6	2	2		3	10					
7				1			29	2		
8					1		47	3		
9									77	

Рис. 7а. Многослойный персептрон, обучающая выборка

Fig. 7a. A multi-layer perceptron, training sample

True class \ Predicted class	1	2	3	4	5	6	7	8	9
1	17								
2		15							
3		1	50		2			2	
4			2	1	2		1	1	
5	2		1		31	1		1	
6					5	2			
7				3			9	2	
8					1	2	15	1	
9									33

Рис. 7б. Многослойный персептрон, тестирующая выборка

Fig. 7b. A multi-layer perceptron, training sample

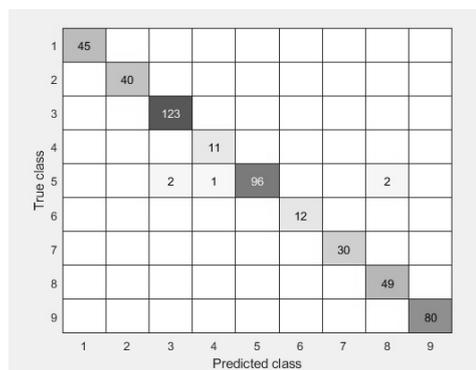


Рис. 7в. Сеть распознавания, обучающая выборка

Fig. 7c. A recognition network, a training sample

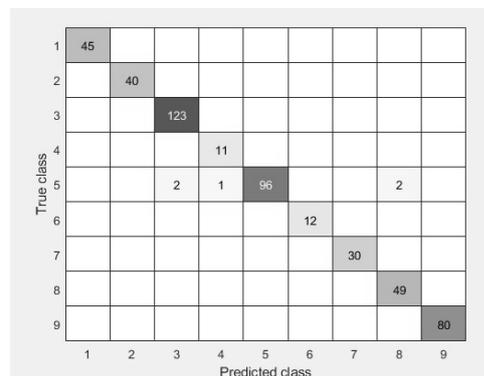


Рис. 7г. Сеть распознавания, тестирующая выборка

Fig. 7d. A recognition network, a testing sample

Кроме того, получены средние значения качества соответственно для тех же случаев, а именно: качество обучающей выборки Q_{train} вычисляется по формуле

$$AVG = \sum_{i=1}^9 \frac{\max_i}{9}, \quad (6)$$

где \max_i – максимальный процент совпадения фактической принадлежности моделей станков и результатов классификации по нейронной сети к i -му классу для обучающей и тестирующей выборки соответственно.

В результате были получены следующие значения:

$AVG_{MP_{train}} = 74.9$, $AVG_{MP_{test}} = 70.4$ – обучающая и тестирующая выборка для многослойного персептрона;

$AVG_{RN_{train}} = 95.3$, $AVG_{RN_{test}} = 89.8$ – обучающая и тестирующая выборка для сети распознавания.

Таким образом, можно сделать вывод, что сеть распознавания дает лучшее качество для выборки моделей станков, чем многослойный персептрон.

ЗАКЛЮЧЕНИЕ

Предложенная в данной работе методика позволяет обучать нейронные сети, которые смогут анализировать выборки станков и определять их принадлежность к определенной группе. В работе для анализа выборки станков применены нейронные сети типов «многослойный персептрон» и «сеть распознавания». Методом главных компонент (РСА) было сокращено пространство признаков, что позволило алгоритму более рационально и быстро выполнять операции.

Полученные результаты позволят специалистам, занимающимся анализом характеристик станков, при покупке новой модели принимать ре-

шение о ее принадлежности определенному типу более быстро и рационально.

Эта методика эффективна именно для рассматриваемой выборки моделей станков и, возможно, будет иной для другой предметной области.

СПИСОК ЛИТЕРАТУРЫ

1. Горбань А.Н. Обучение нейронных сетей. – М.: Параграф, 1990. – 160 с.
2. Нейроинформатика / А.Н. Горбань и др.; отв. ред. Е.А. Новиков. – Новосибирск: Наука, 1998. – 296 с.
3. Гудфеллоу Я., Бенджио И., Курвилль А. Глубокое обучение. – М.: ДМК-Пресс, 2017. – 652 с.
4. Каллан Р. Основные концепции нейронных сетей. – М.: Вильямс, 2001. – 288 с.
5. Круглов В.В., Борисов В.В. Искусственные нейронные сети: теория и практика. – М.: Горячая линия-Телеком, 2001. – 382 с.
6. Goodfellow Y., Benjio I., Courville A. Deep learning. – Cambridge: The MIT Press, 2016.
7. Merrel J.-P. Neural network. – Dordrecht: Springer, 2006. – P. 15–79.
8. Menezes A.J., Oorschot P.V., Vanstone S.A. Handbook of applied cryptography. – Boca Raton: CRC Press, 1996. – 816 p.
9. Терехов В.А., Ефимов Д.В., Тюкин И.Ю. Нейросетевые системы управления. – М.: Высшая школа, 2002. – 184 с.

Рогачева Софья Андреевна, аспирант кафедры вычислительной техники Иркутского национального исследовательского технического университета. Основное направление научных исследований – машинное обучение, нейронные сети, искусственный интеллект. E-mail: pyatsot@bk.ru

Сосинская Софья Семеновна, кандидат технических наук, профессор кафедры вычислительной техники Иркутского национального исследовательского технического университета. Основное направление научных исследований – искусственный интеллект, распознавание образов, параллельные вычисления. Имеет более 70 печатных работ и учебных пособий. E-mail: sosinskaya@mail.ru

Rogacheva Sofya A., postgraduate student at the department of computer engineering, Irkutsk National Research Technical University. Her basic research areas include machine learning, neural networks, and artificial intelligence. E-mail: pyatsot@bk.ru

Sosinskaya Sofya S., PhD (Eng.), professor, department of computing engineering, INRTU. Her research interests are currently focused on artificial intelligence, pattern recognition, and parallel computing. She has more than 70 publications and teaching manuals. E-mail: sosinskaya@mail.ru

DOI: 10.17212/1814-1196-2020-1-107-118

Classification of machine tools described by a large number of features based on the neural network of recognition and attribute space reduction *S.S. SOSINSKAYA^a, S.A. ROGACHEVA^b*Irkutsk National Research Technical University, 83 Lermontov Street, Irkutsk, 664074, Russian Federation*^a *sosinskaya@mail.ru* ^b *pyatsot@bk.ru***Abstract**

In recent years, the relevance of classifying information using neural networks has been growing. This is due to the fact that the amount of data that needs to be processed is getting larger every day. Almost all modern software systems are characterized by a large variety of interacting software modules, which, in turn, increases the complexity of data processing. At this stage, software developers use different methods. Neural networks allow you to achieve the highest possible speeds and high accuracy in working with a large amount of information compared to other methods of classification and data processing. In this regard, the task of developing classification methods is becoming more and more urgent.

The article considers an approach to classifying a sample of machine models described by a large number of features. Two of the most well-known types of neural networks are used as a classification tool, namely the multi-layer perceptron and the recognition network. Due to the fact that no high quality classification was obtained on any network type using a full set of features, the authors applied the principal component method (PCA) to reduce the feature space, which resulted in a significant increase in the quality of classification.

The developed approach can be used to classify machine models that are not represented in the sample. In addition, the article illustrates the fact that the choice of a classification method largely depends on the type of a subject area and the nature of the sample.

The software implementation is based on the MATLAB system, which provides a variety of tools and methods for preparing, analyzing, and visualizing data.

Keywords: classification, neural networks, method development, multi-layer perceptron, recognition network, principal component method, RSA method, feature space reduction

REFERENCES

1. Gorban' A.N. *Obuchenie neironnykh setei* [Training of neural networks]. Moscow, Paragraph Publ., 1990. 160 p. (In Russian).
2. Gorban' A.N., Dunin-Barkovskii V.L. e. a. *Neuroinformatika* [Neuroinformatics]. Novosibirsk, Nauka Publ., 1998. 296 p.
3. Goodfellow Y., Benjio I., Courville A. *Deep learning*. Cambridge, The MIT Press, 2017 (Russ. ed.: Gudfellou Ya., Bendzhio I., Kurvill' A. *Glubokoe obuchenie*. Moscow, DMK-Press, 2017. 652 p.).
4. Callan R. *The essence of neural networks*. London, Prentice Hall Europe, 1999 (Russ. ed.: Kallan R. *Osnovnye kontseptsii neironnykh setei*. Moscow, Williams Publ., 2001. 288 p.).
5. Kruglov V.V., Borisov V.V. *Iskusstvennye neironnye seti: teoriya i praktika* [Artificial neural networks. Theory and practice]. Moscow, Hotline-Telecom, 2001. 382 p.
6. Goodfellow Y., Benjio I., Courville A. *Deep learning*. Cambridge, The MIT Press, 2016.
7. Merrel J.-P. *Neural network*. Dordrecht: Springer, 2006. P. 15–79.
8. Menezes A.J., Oorschot P.V., Vanstone S.A. *Handbook of applied cryptography*. Boca Raton, CRC Press, 1996. 816 p.

* Received 30 October 2019.

9. Terekhov V.A., Efimov D.V., Tyukin I.Yu. *Neurosetevye sistemy upravleniya* [Neural network control system]. Moscow, Vysshaya shkola Publ., 2002. 184 p.

Для цитирования:

Сосинская С.С., Рогачева С.А. Классификация станков, описываемых большим количеством признаков, на основе нейронной сети распознавания и редукции пространства признаков // Научный вестник НГТУ. – 2020. – № 1 (78). – С. 107–118. – DOI: 10.17212/1814-1196-2020-1-107-118.

For citation:

Sosinskaya S.S., Rogacheva S.A. Klassifikatsiya stankov, opisyyvaemykh bol'shim kolichestvom priznakov, na osnove neironnoi seti raspoznavaniya i reduksii prostranstva priznakov [Classification of machine tools described by a large number of features based on the neural network of recognition and attribute space reduction]. *Nauchnyi vestnik Novosibirskogo gosudarstvennogo tekhnicheskogo universiteta = Science bulletin of the Novosibirsk state technical university*, 2020, no. 1 (78), pp. 107–118. DOI: 10.17212/1814-1196-2020-1-107-118.