Studies in Computational Intelligence

Volume 661

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

About this Series

The series "Studies in Computational Intelligence" (SCI) publishes new developments and advances in the various areas of computational intelligence—quickly and with a high quality. The intent is to cover the theory, applications, and design methods of computational intelligence, as embedded in the fields of engineering, computer science, physics and life sciences, as well as the methodologies behind them. The series contains monographs, lecture notes and edited volumes in computational intelligence spanning the areas of neural networks, connectionist systems, genetic algorithms, evolutionary computation, artificial intelligence, cellular automata, self-organizing systems, soft computing, fuzzy systems, and hybrid intelligent systems. Of particular value to both the contributors and the readership are the short publication timeframe and the worldwide distribution, which enable both wide and rapid dissemination of research output.

More information about this series at http://www.springer.com/series/7092

Iuliana F. Iatan

Issues in the Use of Neural Networks in Information Retrieval

Springer

Iuliana F. Iatan
Department of Mathematics and Computer Sciences
Technical University of Civil Engineering
Bucharest
Romania

ISSN 1860-949X ISSN 1860-9503 (electronic)
Studies in Computational Intelligence
ISBN 978-3-319-43870-2 ISBN 978-3-319-43871-9 (eBook)
DOI 10.1007/978-3-319-43871-9

Library of Congress Control Number: 2016948299

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

The measure of success for a person is the magnitude of his/her ability to convert negative conditions to positive ones and achieve goals.

—G.A. Anastassiou

Contents

Introduction

The pattern recognition field is a fundamental one and it is in a full expansion direction of the information technology. The pattern search in the data has a long and a full success history. The observation of some regular patterns in the planet motion has determined Kepler to the discovery of the empirical laws of the planetary motion. The discovery of some regularities in the atomic spectrum has played a key role in the quantum physical development.

In the last years, pattern recognition (PR) has some essential applications in the biometry, satellite image analysis for the detection and the assessment of the terrestrial resources, robotics, medicine, biology, psychology, marketing, computer vision, artificial intelligence, and remote sensing. Other effervescent fields have detached from the PR field: data mining, Web searching, retrieval of multimedia data. Recently, "a lot of area comes under Pattern Recognition due to emerging application which are not only challenging but also computationally more demanding"[1] (see Fig. 1).

"Pattern recognition is not only about methods; it is about taking a new view of the problem at hand that allows one to single out the principal point of interest in a large volume of information and suggest a nontrivial solution."[2]

The statistical approach has been most intensively studied and used in practice than the traditional approaches of pattern recognition. However, the theory of artificial neural network techniques has been getting significant importance. "The design of a recognition system requires careful attention to the following issues:

[1]Dutt, V., and Chadhury, V., and Khan, I., Different Approaches in Pattern Recognition, Computer Science and Engineering, 2011, 1(2), 32–35.

[2]Neimark, Yu.I., and Teklina, L.G., On Possibilities of Using Pattern Recognition Methods to Study Mathematical Models, Pattern Recognition and Image Analysis, 2012, 22(1), 144–149.

Problem Domain	Application	Input Pattern	Pattern classes
Bioinformatics	Sequence Analysis	DNA/ Protein Sequence	Known types of genes patterns
Data Mining	Searching for meaningful patterns	Points in multi- dimensional space	Compact and well separated clusters
Document image analysis	Reading machine for the blind	Document image	Alphanumeric characters, words
Multimedia database retrieval	Internet search	Video clip	Video genres(e.g. action, dialogue,etc.)
Biometric recognition	Forcasting crop yield	Multispectral image	Land use categories, growth pattern of crops
Speech recognition	Telephone directory enquiry without operator assistance	Speech waveform	Spoken words

Fig. 1 Example of pattern recognition applications

definition of pattern classes, sensing environment, pattern representation, feature extraction and selection, cluster analysis, classifier design and learning, selection of training and test samples, and performance evaluation."[3]

Based on presence or absence of teacher and the information provided for the system to learn there are three basic types of learning methods in neural networks [1]:

1. Supervised learning;
2. Unsupervised learning;
3. Reinforced learning;

These are further categorized, based on the rules used, as follows:

- Hebbian,
- Gradient descent,
- Competitive,
- Stochastic learning.

[3]Basu, J.K., and Bhattacharyya, D., and Kim, T.H., Use of Artificial Neural Network in Pattern Recognition, International Journal of Software Engineering and Its Applications, 2010, 4(2), 22–34.

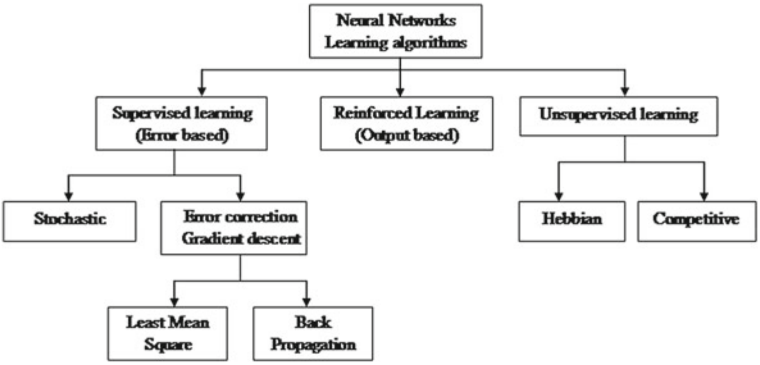

Fig. 2 Classification of learning algorithms

Figure 2 indicates the hierarchical representation of the previous mentioned algorithms.

The four best known approaches for pattern recognition are as follows:

1. Statistical classification
2. Syntactic or structural matching
3. The recognition with intelligence computational techniques
4. The recognition with the expert systems based on some languages specific to the artificial intelligence.

Neural computing is an information processing paradigm, inspired by biological system, composed [2] of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems.

Neural network constitutes an important component in Artificial Intelligence (AI). "It has been studied for many years in the hope of achieving human-like performance in many fields, such as classification, clustering, and pattern recognition, speech and image recognition as well as information retrieval by modeling the human neural system."[4]

Artificial neural network (ANN) represents an adaptive mathematical model or a computational structure, designed to simulate a system of biological neurons, which transfers an information from its input to output in a desired way. An artificial neural network has a lot of interconnecting artificial neurons to employ some mathematical or computational models for information processing. Among the advantages of the neural networks are: learning, adaption, fault tolerance, parallelism, and generalization.

For several years now, neural networks (NNs) have been applied in many different application domains in order to solve various information processing

[4]Reshadat, V., and Feizi-Derakhshi, M.R., Neural Network-Based Methods in Information Retrieval, American Journal of Scientific Research, 2011, 58, 33–43.

problems of regression, classification, computational science, computer vision, data processing, and time series analysis. Hence, they have enjoyed wide popularity [3].

Neural network applications can be grouped [1] in the following categories:

(A) **Clustering**: It is a "method to organize automatically a large data collection by partition a set data, so the objects in the same cluster are more similar to one another than with the objects belonging to other clusters."[5] The representative applications include data compression and data mining.

(B) **Classification/Pattern Recognition**: It has the task to assign an input pattern to one of many classes. This category includes algorithmic implementations such as associative memory.

(C) **Function Approximation**: As its aim is to find an estimate of the unknown function subject to noise, it is required by various engineering and scientific disciplines.

(D) **Prediction Systems**: They have the task to forecast some future values of a time-sequenced data. Prediction is unlike the function approximation by considering time factor. Hence, the dynamic system may produce different results for the same input data based on time, which means the system state.

Information retrieval (IR) represents a wide research area mainly on the Internet. IR "is concerned with the analysis, representation and retrieval of texts."[6]

IR "is different from data retrieval in databases using SQL queries because the data in databases are highly structured and stored in relational tables, while information in text is unstructured. There is no structured query language like SQL for text retrieval."[7]

Data mining constitutes another layer of data processing, introduced in order to have a better perception of information for management and decision making. The main aim of this processing layer is [4] to:

(i) extract the implicit, hidden, potentially useful information;
(ii) discover meaning full patterns from large raw data collections.

Data mining is "a multidisciplinary field involving machine learning, statistics, databases, artificial intelligence, information retrieval, and visualization." (see footnote 7).

Some of the common data mining tasks are [5]: supervised learning (or classification), unsupervised learning (or clustering), association rule mining, sequential pattern mining, and regression.

Document clustering is a fundamental task in text mining, being concerned with grouping documents into clusters according to their similarity; more exactly,

---

[5]Bharathi, G. and Venkatesan, D., Improving Information Retrieval using Document Clusters and Semantic Synonym Extraction, Journal of Theoretical and Applied Information Technology, 2012, 36(2), 167–173.

[6]Reshadat, V., and Feizi-Derakhshi, M.R., Neural Network-Based Methods in Information Retrieval, American Journal of Scientific Research, 2011, 58, 33–43.

[7]Liu, B., Web DataMining, Springer-Verlag Berlin Heidelberg, 2008.

document clustering achieves automatically group of the documents that belong to the same topic, in order to provide user's browsing of retrieval results. "Document clustering has always been used as a tool to improve the performance of retrieval and navigating large data."[8] The representative applications include data compression and data mining.

Being suited for information retrieval from large text to multimedia databases, the NNs have been widely used in the area of IR and text mining, such as text classification, text clustering, and collaborative filtering. "In recent years, with the fast growth of the World Wide Web and the Internet, these algorithms have also been used in Web-related applications such as Web searching, Web page clustering, Web mining, etc. Their capacity for tolerant and intuitive processing offers new perspectives in information retrieval."(see footnote 6).

When large volumes of data are to be handled, a suitable approach to increase the IR speed is to use the NNs as an artificial intelligent technique. I was motivated to write this book to highlight the ability of the NNs to be very good pattern matchers and the importance of the NNs for the IR, which is based on index term matching.

Chapter 1 of this book emphasizes the possibility of using neural networks in IR and highlights the advantages of applying two neural networks models for solving the problem of simplifying the complex structure of an IR system, by substitution of the relations between its subsystems by NNs.

The core to measure similarity or distance between two information entities is required for all information discovery tasks (whether IR or data mining). It is crucial to use an appropriate measure both to improve the quality of information selection and to reduce the time and processing costs [4]. The relevance of the concept of similarity has proven [4] not only in every scientific field but in philosophy and psychology, too. However, this work deals with the measure of similarity in computer science domain (information retrieval and data mining to be more specific). In the thesis domain, the similarity measure "is an algorithm that determines the degree of agreement between entities."[9]

"Similarity-based classifiers estimate the class label of a test sample based on the similarities between the test sample and a set of labeled training samples, and the pairwise similarities between the training samples."[10]

Image similarity is an important concept in many applications. In Chap. 2 of this book we define a new neural network based method for learning image similarity [6]. We start off from the fuzzy Kwan-Cai neural network [7] and turn it into a partially supervised one. Using the training algorithm of the FKCNN, its third and

---

[8]Bharathi, G. and Venkatesan, D., Improving Information Retrieval using Document Clusters and Semantic Synonym Extraction, Journal of Theoretical and Applied Information Technology, 2012, 36(2), 167–173.

[9]Zaka, B., Theory and Applications of Similarity Detection Techniques, 2009, http://www.iicm.tugraz.at/thesis/bilal_dissertation.pdf.

[10]Chen, Y., and Garcia, E.K., and Gupta, M.Y., and Rahimi, A., and Cazzanti, A., Similarity-based Classification: Concepts and Algorithms, Journal of Machine Learning Research, 2009, 10, 747–776.

fourth layers are built during the learning process. The training stage is followed by a calibration stage, in which the fuzzy neurons (FNs) of the fourth layer will be assigned category labels. We performed a comparative studiy of the proposed similarity learning method and compared it to self-organizing Kohonen maps (SOKM) and *k*-nearest neighbor rule (*k*-NN). The resulting similarity functions are tested on the VOC data set that consists in 20 object classes. The results indicate that the neural methods FKCNN and SOKM are performing better for our task than *k*-NN. SOKM sometimes gives good results, but this depends highly on the right parameter settings. Small variations induced large drops in performance. The overall performance of FKCNN is better. The main advantage of FKCNN consists in the fact that we can obtain good results that are robust to changes in the parameter settings.

There is a section in Chap. 2, where we have performed the software implementation of the FKCNN to be experimented for the face recognition task, using the ORL Database of Faces, provided by the AT&T Laboratories from Cambridge University; it contains 400 images, corresponding to 40 subjects (namely, 10 images for each of the 40 classes).

Recently, ANNs methods have become useful for a wide variety of applications across a lot of disciplines and in particularly for prediction, where highly nonlinear approaches are required [8]. The advantage of neural networks consists [9] in their ability to represent both linear and nonlinear relationships and to learn these relationships directly from the data being modeled. Among the statistical techniques that are widely used is the regression method, the multiple regression analysis has the objective to use independent variables whose values are known to predict the single dependent variable.

Our research objective in Chap. 3 is to compare [10] the predictive ability of multiple regression and fuzzy neural model, by which a user's personality can be accurately predicted through the publicly available information on their Facebook profile. We shall choose to use the fuzzy Gaussian neural network (FGNN) for predicting personality because it handles nonlinearity associated with the data well.

Function approximation has the aim to find the underlying relationship from a given finite input–output data, being a fundamental problem in a lot of real-world applications, such as prediction, pattern recognition, data mining, and classification. Various methods have been developed to solve the problem of function approximation, one of them being with artificial neural networks [11]. The problem of estimating a function from a set of samples means an advance in the field of neural networks, as there has been much research work being carried out in exploring the function approximation capabilities of NN's [12]. "Approximation or representation capabilities of neural networks and fuzzy systems have attracted considerable research in the last 15 years."[11]

[11]Zeng, X.J., and Keane, J.A., and Goulermas, J.Y., and Liatsis, P., Approximation Capabilities of Hierarchical Neural-Fuzzy Systems for Function Approximation on Discrete Spaces, International Journal of Computational Intelligence Research, 2005, 1, 29–41.

The Fourier series neural networks (FSNNs) represent one type of orthogonal neural networks (ONN) and they are feedforward networks, similar to sigmoidal neural networks. After we have studied the FSNN for function approximation in Chap. 4, we have designed a new neural model [13]. FSNN performs only to approximate trigonometric functions and not for all the kind of functions. Therefore, it was necessary to built in the paper [13] a four layer neural network which works very well both for the approximation of the trigonometric functions and for other types of functions (as multivariate polynomial, exponential) too. The main advantage of our proposed model consists in its special weight matrix. This matrix has different expressions in terms of the function that has to be approximated.

For each function, the approximation achieved using our neural method is finer than that obtained using the FSNN. Approximating a function with an FSNN is better than using a Tayor series as it has a smaller maximum error and is more economical. There are many advantages of using neural network to implement function approximation.

The field studied in this chapter has significant references for the research of function approximation and all the conclusions are proved effective after the actual simulation tests.

The idea of Chap. 5 is to develop the neural networks in other than the real domain. "Neural computation in Clifford algebras, which include familiar complex numbers and quaternions as special cases, has recently become an active research field."[12] It is interesting to use complex numbers in the context of neural networks as they tend to improve learning ability [14]. "Though neurons with plural real or complex numbers may be used to represent multidimensional data in neural networks, the direct encoding in terms of hyper-complex numbers may be more efficient."[13]

In order to achieve the aim of illustrating the usefulness of the Clifford algebra in the neural computing because of its geometric properties, we have introduced in Chap. 5, the Fuzzy Clifford Gaussian network (FCGNN), contributing [15] in this way to continue the development of neural networks in other than the real domain.

Chapter 6 introduces two concurrent neural models:

1. *Concurrent fuzzy nonlinear perceptron modules* (CFNPM), representing a winner-takes-all collection of small FNP (Fuzzy Nonlinear Perceptron) units;
2. *Concurrent fuzzy Gaussian neural network modules* (CFGNNM), which consists of a set of M fuzzy neural networks, by the type FGNN.

[12] Buchholz, S., and Sommer, G., Introduction to Neural Computation in Clifford Algebra, 2010, http://www.informatik.uni-kiel.de/inf/Sommer/doc/Publications/geocom/buchholz_sommer1.pdf.

[13] Matsui, N., and Isokawa, T., and Kusamichi, H., and Peper, F., and Nishimura, H., Quaternion neural network with geometrical operators, Journal of Intelligent & Fuzzy Systems, 2004, 15, 149–164.

The use of them for face recognition causes an increase in the recognition rates for the training and the test lot (both in the case of making the feature selection), compared to those achieved using the simple variations of FNP and FGNN.

The aim of Chap. 7 is to design a new model of fuzzy nonlinear perceptron, based on alpha level sets, entitled as fuzzy nonlinear perceptron based on alpha level sets (FNPALS). It differs from the other fuzzy variants of the nonlinear perceptron, where the fuzzy numbers are represented by membership values, i.e., in the case of the FNPALS, the fuzzy numbers are represented through the alpha level sets.

The last chapter has the aim to describe a recurrent fuzzy neural network (RFNN) model, whose learning algorithm is based on the improved particle swarm optimization (IPSO) method. The proposed RFNN is different from other variants of RFNN models through the number of the evolution directions that they use: we update the velocity and the position of all particles along three dimensions [16], while in [17] two dimensions are used.

References

1. R.C. Chakraborty. Fundamentals of neural networks. http://www.myreaders.info/html/artificial_intelligence.htm, 2010.
2. J.K. Basu, D. Bhattacharyya, and T.H. Kim. Use of artificial neural network in pattern recognition. *International Journal of Software Engineering and Its Applications*, 4(2):22–34, 2010.
3. L.A. Gougam, M. Tribeche, and F. Mekideche-Chafa. A systematic investigation of a neural network for function approximation. *Neural Networks*, 21:1311–1317, 2008.
4. B. Zaka. Theory and applications of similarity detection techniques. http://www.iicm.tugraz.at/thesis/bilal_dissertation.pdf, 2009.
5. B. Liu. *Web DataMining*. Springer-Verlag Berlin Heidelberg, 2008.
6. I. Iatan and M. Worring. A fuzzy Kwan- Cai neural network for determining image similarity. *BioSystems (Under Review),* 2016.
7. H.K. Kwan and Y. Cai. A fuzzy neural network and its application to pattern recognition. *IEEE Trans. on Fuzzy Systems*, 2(3):185–193, 1997.
8. V. Bourdès, S. Bonnevay, P. Lisboa, R. Defrance, D. Pérol, S. Chabaud, T. Bachelot, T. Gargi, and S. Nègrier. Comparison of artificial neural network with logistic regression as classification models for variable selection for prediction of breast cancer patient outcomes. *Advances in Artificial Neural Systems*, pages 1–10, 2010.
9. O.S. Maliki, A.O. Agbo, A.O. Maliki, L.M. Ibeh, and C.O. Agwu. Comparison of regression model and artificial neural network model for the prediction of electrical power generated in Nigeria. *Advances in Applied Science Research*, 2(5):329–339, 2011.
10. I. Iatan and M. de Rijke. Predicting human personality from social media using a fuzzy neural network. *Neural Computing and Applications (Under Review)*, 2016.
11. Z. Zainuddin and P. Ong. Function approximation using artificial neural networks. *WSEAS Transactions on Mathematics*, 7(6):333–338, 2008.
12. S.S. Haider and X.J. Zeng. Simplified neural networks algorithm for function approximationon discrete input spaces in high dimension-limited sample applications. *Neurocomputing*, 72:1078–1083, 2009.
13. I. Iatan, S. Migorski, and M. de Rijke. Modern neural methods for function approximation. *Applied Soft Computing*, (Under Review), 2016.

14. N. Matsui, T. Isokawa, H. Kusamichi, F. Peper, and H. Nishimura. Quaternion neural network with geometrical operators. *Journal of Intelligent & Fuzzy Systems*, 15:149–164, 2004.
15. I. Iatan and M. de Rijke. A fuzzy Gaussian Clifford neural network. (work in progress), 2014.
16. G.A. Anastassiou and I. Iatan. A recurrent neural fuzzy network. *Journal of Computational Analysis and Applications*, 20 (2), 2016.
17. C.J. Lin, M. Wang, and C.Y. Lee. Pattern recognition using neural-fuzzy networks based on improved particle swam optimization. *Expert Systems with Applications*, 36:5402–5410, 2009.