

## Chapter 11

# DE–Based RBFNs for Classification With Special Attention to Noise Removal and Irrelevant Features

**Ch. Sanjeev Kumar Dash**  
*Silicon Institute of Technology, India*

**Ajit Kumar Behera**  
*Silicon Institute of Technology, India*

**Sarat Chandra Nayak**  
*Kommuri Pratap Reddy Institute of Technology, India*

### ABSTRACT

*This chapter presents a novel approach for classification of dataset by suitably tuning the parameters of radial basis function networks with an additional cost of feature selection. Inputting optimal and relevant set of features to a radial basis function may greatly enhance the network efficiency (in terms of accuracy) at the same time compact its size. In this chapter, the authors use information gain theory (a kind of filter approach) for reducing the features and differential evolution for tuning center and spread of radial basis functions. Different feature selection methods, handling missing values and removal of inconsistency to improve the classification accuracy of the proposed model are emphasized. The proposed approach is validated with a few benchmarking highly skewed and balanced dataset retrieved from University of California, Irvine (UCI) repository. The experimental study is encouraging to pursue further extensive research in highly skewed data.*

DOI: 10.4018/978-1-5225-2857-9.ch011

## 1. INTRODUCTION

Classification is one of the fundamental tasks in data mining and pattern recognition. Over the years many models have been proposed. However, it is a consensus that the accuracy of the discovered model (i.e., a neural networks (NNs), rules, decision tree strongly depends on the quality of the data being mined. Hence inconsistency removal and feature selection brings lots of attention of many researchers. If the inconsistent data is simply deleted or classified as a new category then inevitably some useful information will be lost. The method used in this article for making the dataset consistent is based on Bayesian statistical method. Here the inconsistent data is classified as the most probable one and the redundant data records are deleted as well. So the loss of information due to simple deletion or random classification of inconsistent data is reduced and the size of the dataset is also reduced.

Feature selection is the process of selecting a subset of available features to use in empirical modeling. Like feature selection, instance selection is to choose a subset samples to achieve the original purpose of a classification tasks, as if the whole dataset is used. Many variants of evolutionary and non-evolutionary based approaches are discussed in literatures. The ideal outcome of instance selection is a model independent, minimum sample of data that can accomplish tasks with little or no performance deterioration. Unlike feature selection and instance selection, feature extraction at feature level fusion recently attracts data mining/machine learning researchers to give special focus while designing a classifier.

Feature selection can be broadly classified into two categories: i) filter approach (it depends on generic statistical measurement); and ii) wrapper approach (based on the accuracy of a specific classifier). In this work, the feature selection is performed based on information gain theory (entropy) measure with a goal to select a subset of features that preserves as much as possible the relevant information found in the entire set of features. After selection of the relevant set of features the fine tuned radial basis function network is modeled using differential evolution for classification of both balanced and unbalanced datasets. In imbalance classification problems occurs that the number of instances of each class can be very different.

Over the decade radial basis function (RBF) networks have attracted a lot of interest in various domain of interest. One reason is that they form a unifying link between function approximation, regularization, noisy interpolation, classification, and density estimation. It is also the case that training RBF networks is usually faster than training multi-layer perceptron networks. RBF network training usually proceeds in two steps: First, the basis function parameters (corresponding to hidden units) are determined by clustering. Second, the final-layer weights are determined by least square which reduces to solve a simple linear system. Thus, the first stage is an unsupervised method which is relatively fast, and the second stage requires the solution of a linear problem, which is also fast.

One of the advantages of RBF neural networks, compared to multi-layer perceptron networks, is the possibility of choosing suitable parameters for the units of hidden layer without having to perform a nonlinear optimization of the network parameters. However, the problem of selecting the appropriate number of basis functions remains a critical issue for RBF networks. The number of basis functions controls the complexity, and hence the generalization ability of RBF networks. An RBF network with too few basis functions gives poor predictions on new data, i.e., poor generalization, since the model has limited flexibility. On the other hand, an RBF network with too many basis functions also yields poor generalization since it is too flexible and fits the noise in the training data. A small number of basis functions yields a high bias, low variance estimator, whereas a large number of basis functions yields a low bias but high variance estimator. The best generalization performance is obtained via a compromise

24 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:  
[www.igi-global.com/chapter/de-based-rbfns-for-classification-with-special-attention-to-noise-removal-and-irrelevant-features/187688](http://www.igi-global.com/chapter/de-based-rbfns-for-classification-with-special-attention-to-noise-removal-and-irrelevant-features/187688)

## Related Content

---

### Faster Self-Organizing Fuzzy Neural Network Training and Improved Autonomy with Time-Delayed Synapses for Locally Recurrent Learning

Damien Coyle, Girijesh Prasad and Martin McGinnity (2011). *System and Circuit Design for Biologically-Inspired Intelligent Learning* (pp. 156-183).

[www.irma-international.org/chapter/faster-self-organizing-fuzzy-neural/48895](http://www.irma-international.org/chapter/faster-self-organizing-fuzzy-neural/48895)

### Optimality-Oriented Stabilization for Recurrent Neural Networks

Ziqian Liu (2011). *System and Circuit Design for Biologically-Inspired Intelligent Learning* (pp. 93-115).

[www.irma-international.org/chapter/optimality-oriented-stabilization-recurrent-neural/48892](http://www.irma-international.org/chapter/optimality-oriented-stabilization-recurrent-neural/48892)

### Optimisation of Engineering Systems With the Bees Algorithm

Duc T. Pham, Luca Baronti, Biao Zhang and Marco Castellani (2018). *International Journal of Artificial Life Research* (pp. 1-15).

[www.irma-international.org/article/optimisation-of-engineering-systems-with-the-bees-algorithm/211932](http://www.irma-international.org/article/optimisation-of-engineering-systems-with-the-bees-algorithm/211932)

### City Group Optimization: An Optimizer for Continuous Problems

Yijun Yang (2018). *Critical Developments and Applications of Swarm Intelligence* (pp. 73-96).

[www.irma-international.org/chapter/city-group-optimization/198922](http://www.irma-international.org/chapter/city-group-optimization/198922)

### Sticker P System

A. S. Prasanna Venkatesan, V. Masilamani and D. G. Thomas (2012). *International Journal of Natural Computing Research* (pp. 28-43).

[www.irma-international.org/article/sticker-system/72870](http://www.irma-international.org/article/sticker-system/72870)