Evolutionary q-Gaussian Radial Basis Functions for Binary-Classification

F. Fernández-Navarro, C. Hervás-Martínez,
Pedro A. Gutiérrez, M. Cruz-Ramírez and M. Carbonero-Ruz
pagutierrez@uco.es

Department of Computer Science and Numerical Analysis
AYRNA Research Group
University of Córdoba, Spain.

HAIS'10
5th International Conference on HYBRID ARTIFICIAL INTELLIGENCE SYSTEMS

23rd June, 2010

1. Outline
2. Introduction
3. RBFNNs
4. Evolutionary q-Gaussian RBFs for Classification
5. Experiments
   - Experimental Design
   - Statistical Results
6. Conclusions

Introduction

Different types of neural networks used for classification purposes.

- Projection basis functions:
  - Multilayer perceptron neural networks (MLP).
  - Product Unit Neural Networks (PUNNs).
- Kernel basis functions:
  - Radial Basis Function (RBF) neural networks.

Shortcommings of the standard RBF

When dimensionality grows and/or when data is concentrated in boundaries of the K dimensional space, standard Gaussian basis function lacks its performance.

Our Proposal

Alternative q-Gaussian Radial Basis Neural Network obtained by a Hybrid Algorithm (HA).

Radial Basis Functions Neural Networks

The model of a RBFNN can be described with the following equation (binary classification or regression):

$$f(\mathbf{x}) = \beta_0 + \sum_{i=1}^{m} \beta_i \cdot \phi_i(d_i(\mathbf{x})) \quad (1)$$

The function $d_i(\mathbf{x})$ is defined as:

$$d_i(\mathbf{x}) = \frac{\|\mathbf{x} - \mathbf{c}_i\|^2}{r_i^2}, 1 \leq i \leq m \quad (2)$$

Radial Basis Functions Neural Networks

Standard RBF (SRBF)

$$\phi_i(d_i(\mathbf{x})) = e^{-d_i(\mathbf{x})} \quad (3)$$

- Very selective response, with high activation for patterns close to the centroid and very small activation for distant patterns.

Radial Basis Functions Neural Networks

Cauchy RBF (CRBF)

$$\phi_i(d_i(\mathbf{x})) = \frac{1}{1 + d_i(\mathbf{x})} \quad (4)$$

Inverse Multiquadratic RBF (IMRBF)

$$\phi_i(d_i(\mathbf{x})) = \frac{1}{(1 + d_i(\mathbf{x}))^{\frac{1}{2}}} \quad (5)$$

- The CRBF and IMRBF have longer tails than the SRBF. $\rightarrow$ Activation for patterns distant to the centroid of the RBF is bigger than the activation of the SRBF for those patterns.

q-Gaussian RBFs

q-Gaussian RBF

$$\phi_i(d_i(\mathbf{x})) = \begin{cases} (1-(1-q)d_i(\mathbf{x}))^{\frac{1}{1-q}} & \text{if } (1-(1-q)d_i(\mathbf{x})) \geq 0; \\ 0 & \text{Otherwise} \end{cases} \quad (6)$$

- The q-Gaussian can reproduce different RBFs for different values of the real parameter q.
  - $q \rightarrow 2$; q-Gaussian≡CRBF.
  - $q \rightarrow 3$; q-Gaussian≡IMRBF.
  - $q \rightarrow 1$; q-Gaussian≡SRBF.
- A small change in the value of q represents a smooth modification on the shape of the RBF.

q-Gaussian RBFs

Radial unit activation in one-dimensional space with $c = 0$ and $r = 1$ for different RBFs: (a) Gaussian, Cauchy and Inverse Multiquadratic and (b) q-Gaussian with different values of q

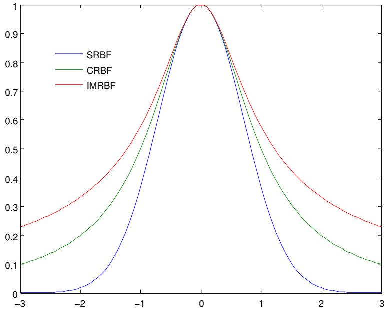

Figure: (a) Alternative RBFs

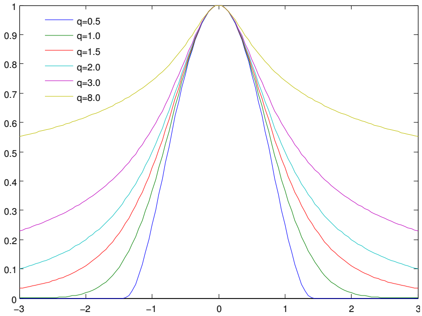

Figure: (b) q-Gaussian RBF

q-Gaussian RBFs for binary classification

- Probabilistic framework:
  - The activation function of each output node is the softmax function.

$$g(\mathbf{x}) = \frac{\exp f(\mathbf{x})}{1 + \exp f(\mathbf{x})} \quad (7)$$

  - It is possible to evaluate the model using the cross-entropy error function, given by:

$$l(g) = -\frac{1}{N}\sum_{n=1}^{N}[y_n \log f(\mathbf{x}_n) + (1 - y_n)\log(1 - f(\mathbf{x}_n))] \quad (8)$$

Hybrid Algorithm

```
Hybrid Algorithm:
Generate a random population of size N
repeat
    Calculate the fitness of every individual in the population
    Rank the individuals with respect to their fitness
    The best individual is copied into the new population
    The best 10% of population individuals are replicated and they
    substitute the worst 10% of individuals
    Apply parametric mutation to the best (p_m)% of individuals
    Apply structural mutation to the remaining (100 − p_m)% of
    individuals
until the stopping criterion is fulfilled
Apply iRprop+ to the best solution obtained by the EA in the last
generation.
```

Figure: Hybrid Algorithm (HA) framework

Hybrid Algorithm (Main characteristics I)

- Error and Fitness Functions.
  - $l(g)$ as the error function.
  - $A(g) = \frac{1}{1+l(g)}$, where $0 < A(g) \leq 1$ as the fitness measure.
- Initialization of the Population. $5,000$ random RBFNNs:
  - k-means algorithm for different values of k, $k \in [M_{min}, M_{max}]$, where M_{min} and M_{max} are parameters of the algorithm.
  - Widths (r_i) of the RBFNNs $\rightarrow$ geometric mean of the distance to the two nearest neighbours.
  - q_i $\rightarrow$ values close to 1 (SRBF).
  - Then we select the best 500 RBFNNs, and we evolve them.

Hybrid Algorithm (Main characteristics II)

- Structural Mutation. There are four different structural mutations: hidden node addition, hidden node deletion, connection addition and connection deletion. If the structural mutator adds a new node in the RBFNN, the q parameter is assigned to a value in the interval $[0.75, 1.25]$.
- Parametric Mutation. Centre, Radii and q Mutation and Output-to-Hidden Node Connection Mutations $\rightarrow$ adding a Gaussian noise.
- iRprop+. We have carried out the adaptation of the *iRprop+* local improvement procedure to the softmax activation function and the cross-entropy error function.

Experimental Design

- Eleven binary classification datasets taken from the UCI repository.
- The performance of each method has been evaluated using the correct classification rate (C) in the generalization set.
- The experimental design was conducted using a 10-fold cross-validation procedure, with 10 repetitions per each fold.
- Comparison of the results obtained to:
    - SRBF.
    - CRBF.
    - IMRBF.

Statistical Results (Mean±Standard Deviation)

	Method(C_{G}(%))			
	SRBF	CRBF	IMRBF	q-Gaussian
Labor	91.33 ± 12.09	**95.00 ± 11.24**	91.66 ± 8.78	*93.33 ± 11.65*
Promoters	75.54 ± 13.56	80.18 ± 6.66	*81.09 ± 8.69*	**84.00 ± 6.15**
Hepatitis	**86.33 ± 8.09**	83.16 ± 7.15	85.12 ± 7.52	*85.30 ± 7.54*
Sonar	**78.38 ± 9.03**	74.09 ± 10.20	76.02 ± 11.16	*76.04 ± 13.56*
Heart	81.85 ± 8.97	83.70 ± 8.76	**84.81 ± 8.45**	*84.07 ± 7.20*
BreastC	72.04 ± 6.39	71.35 ± 8.00	**73.10 ± 6.39**	*73.06 ± 6.77*
Heart-C	85.44 ± 3.83	85.45 ± 5.59	*85.77 ± 3.05*	**85.79 ± 5.20**
Liver	*68.41 ± 5.15*	65.23 ± 8.23	65.52 ± 6.31	**71.30 ± 6.50**
Vote	**96.32 ± 3.97**	95.39 ± 3.59	94.94 ± 2.36	*96.08 ± 3.45*
Card	86.08 ± 3.14	*86.52 ± 3.55*	85.94 ± 3.80	**87.87 ± 0.37**
German	74.80 ± 3.82	*74.90 ± 3.17*	74.40 ± 2.50	**75.25 ± 2.98**
$\overline{C}_{\mathrm{G}}$(%)	81.50	81.36	*81.67*	**82.91**
$\overline{R}$	*2.72*	2.99	2.72	**1.54**
p-Value	**0.03**	**0.00**	**0.03**	-
α'_{Hommel}	0.10	0.03	0.05	-

The best result is in bold face and the second best result in italics

Statistical Results (Mean±Standard Deviation)

	Method(C_{G}(%))			
	SRBF	CRBF	IMRBF	q-Gaussian
Labor	91.33 ± 12.09	**95.00 ± 11.24**	91.66 ± 8.78	*93.33 ± 11.65*
Promoters	75.54 ± 13.56	80.18 ± 6.66	*81.09 ± 8.69*	**84.00 ± 6.15**
Hepatitis	**86.33 ± 8.09**	83.16 ± 7.15	85.12 ± 7.52	*85.30 ± 7.54*
Sonar	**78.38 ± 9.03**	74.09 ± 10.20	76.02 ± 11.16	*76.04 ± 13.56*
Heart	81.85 ± 8.97	83.70 ± 8.76	**84.81 ± 8.45**	*84.07 ± 7.20*
BreastC	72.04 ± 6.39	71.35 ± 8.00	**73.10 ± 6.39**	*73.06 ± 6.77*
Heart-C	85.44 ± 3.83	85.45 ± 5.59	*85.77 ± 3.05*	**85.79 ± 5.20**
Liver	*68.41 ± 5.15*	65.23 ± 8.23	65.52 ± 6.31	**71.30 ± 6.50**
Vote	**96.32 ± 3.97**	95.39 ± 3.59	94.94 ± 2.36	*96.08 ± 3.45*
Card	86.08 ± 3.14	*86.52 ± 3.55*	85.94 ± 3.80	**87.87 ± 0.37**
German	74.80 ± 3.82	*74.90 ± 3.17*	74.40 ± 2.50	**75.25 ± 2.98**
$\overline{C}_{\mathrm{G}}$(%)	81.50	81.36	*81.67*	**82.91**
$\overline{R}$	*2.72*	2.99	2.72	**1.54**
p-Value	**0.03**	**0.00**	**0.03**	-
α'_{Hommel}	0.10	0.03	0.05	-

The best result is in bold face and the second best result in italics

Statistical Results (Mean±Standard Deviation)

	Method($C_{\mathrm{G}}(\%)$)			
	SRBF	CRBF	IMRBF	q-Gaussian
Labor	91.33 ± 12.09	**95.00 ± 11.24**	91.66 ± 8.78	*93.33 ± 11.65*
Promoters	75.54 ± 13.56	80.18 ± 6.66	*81.09 ± 8.69*	**84.00 ± 6.15**
Hepatitis	**86.33 ± 8.09**	83.16 ± 7.15	85.12 ± 7.52	*85.30 ± 7.54*
Sonar	**78.38 ± 9.03**	74.09 ± 10.20	76.02 ± 11.16	*76.04 ± 13.56*
Heart	81.85 ± 8.97	83.70 ± 8.76	**84.81 ± 8.45**	*84.07 ± 7.20*
BreastC	72.04 ± 6.39	71.35 ± 8.00	**73.10 ± 6.39**	*73.06 ± 6.77*
Heart-C	85.44 ± 3.83	85.45 ± 5.59	*85.77 ± 3.05*	**85.79 ± 5.20**
Liver	*68.41 ± 5.15*	65.23 ± 8.23	65.52 ± 6.31	**71.30 ± 6.50**
Vote	**96.32 ± 3.97**	95.39 ± 3.59	94.94 ± 2.36	*96.08 ± 3.45*
Card	86.08 ± 3.14	*86.52 ± 3.55*	85.94 ± 3.80	**87.87 ± 0.37**
German	74.80 ± 3.82	*74.90 ± 3.17*	74.40 ± 2.50	**75.25 ± 2.98**
$\overline{C}_{\mathrm{G}}(\%)$	81.50	81.36	*81.67*	**82.91**
$\overline{R}$	*2.72*	2.99	2.72	**1.54**
p-Value	**0.03**	**0.00**	**0.03**	-
α'_{Hommel}	0.10	0.03	0.05	-

The best result is in bold face and the second best result in italics

Conclusions

Conclusions

- The models proposed, q-Gaussian Radial Basis Functions as transfer functions, are a viable alternative for obtaining more accurate binary classifications.
- These models have been designed with a HA constructed specifically for taking into account the characteristics of this kernel model.

Future research

- To study other alternative RBFs (Generalized RBFs).
- To consider multi-class problems.

Evolutionary q-Gaussian Radial Basis Functions for Binary-Classification

F. Fernández-Navarro, C. Hervás-Martínez,
Pedro A. Gutiérrez, M. Cruz-Ramírez and M. Carbonero-Ruz
pagutierrez@uco.es

Department of Computer Science and Numerical Analysis
AYRNA Research Group
University of Córdoba, Spain.

HAIS'10
5th International Conference on HYBRID ARTIFICIAL INTELLIGENCE SYSTEMS

23rd June, 2010