Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 103 (2017) 308 – 315

Procedia Computer Science

XIIth International Symposium «Intelligent Systems», INTELS'16, 5-7 October 2016, Moscow, Russia

Application of unnormalized neurostructural models ANFIS for the tasks of binary classification

O.A. Nazarkin*, P.V. Saraev

Lipetsk State Technical University, 30, Moskovskaya str., Lipetsk 398600, Russia

Abstract

In this paper, we consider the construction of binary classifiers ensembles on the basis of unnormalized form of ANFIS models. We notice training performance boost due to simplified model structure. In addition, the unnormalized ANFIS architecture provides the ground for effective task parallel decomposition of training procedures.

Peer-review under responsibility of the scientific committee of the XIIth International Symposium "Intelligent Systems"

Keywords: neurostructural models; data mining and database knowledge discovery; function approximators synthesis; binary classifiers; ensemble learning

1. Definition of unnormalized ANFIS model with Gaussian functions

As discussed in[1], knowledge discovery tasks naturally lead to hybrid neurofuzzy, or neurostructural models[2]. With neurostructural models, such as ANFIS[3], it is possible to extract meaningful knowledge in the form of "IF-THEN" rule sets. Function approximators of this type introduce universal approximation capabilities, combined with the ability to represent knowledge in a human-consumable way. Neurostructural ANFIS approximator[3] with rules count R, input vector $\mathbf{x}$ of dimension M, and scalar output is defined by the tuple of adjustable parameters $p = \{a_{rm}, q_{rm}, k_{rm}, b_r\}$, $r = 1..R$, $m = 1..M$.

Models of this class are commonly used in normalized form (1):

* Corresponding author.
E-mail address: nazarkino@mail.ru

Peer-review under responsibility of the scientific committee of the XIIth International Symposium "Intelligent Systems"
doi:10.1016/j.procs.2017.01.113

$$A_p(\mathbf{x}) = \sum_{r=1}^{R} \frac{\prod_{m=1}^{M} \mu_{rm}(x_m)}{\sum_{j=1}^{R}\prod_{m=1}^{M} \mu_{jm}(x_m)} L_r(\mathbf{x}) = \sum_{r=1}^{R} \frac{e^{-\sum_{m=1}^{M}\left(\frac{x_m - a_{rm}}{q_{rm}}\right)^2}}{\sum_{j=1}^{R} e^{-\sum_{m=1}^{M}\left(\frac{x_m - a_{jm}}{q_{jm}}\right)^2}} \left(b_r + \sum_{m=1}^{M} k_{rm} x_m\right) \quad (1)$$

The component (2) is responsible for normalization:

$$\rho(\mathbf{x}) = \sum_{r=1}^{R}\prod_{m=1}^{M} \mu_{jm}(x_m) = \sum_{r=1}^{R} e^{-\sum_{m=1}^{M}\left(\frac{x_m - a_{jm}}{q_{jm}}\right)^2}, \ \rho(\mathbf{x}) > 0 \quad (2)$$

Models of structurally similar class RBFN (Radial Basis Function Networks) have both normalized and unnormalized forms[4]. Normalized RBFNs may demonstrate better generalization abilities, but their unnormalized counterparts are simpler to calculate, therefore, they require less computational resources for training procedure.

In this work, we consider construction of unnormalized form (3) for neurostructural approximators ANFIS with Gaussian functions:

$$U_p(\mathbf{x}) = \rho(\mathbf{x}) A_p(\mathbf{x}) = \sum_{r=1}^{R}\prod_{m=1}^{M} \mu_{rm}(x_m) L_r(\mathbf{x}) = \sum_{r=1}^{R} e^{-\sum_{m=1}^{M}\left(\frac{x_m - a_{rm}}{q_{rm}}\right)^2} \left(b_r + \sum_{m=1}^{M} k_{rm} x_m\right). \quad (3)$$

The goal of model training is to find optimal values for parameters p_T^* on the training set $T = \{\mathbf{x}_i, f(\mathbf{x}_i)\}$ of dimension N. This task is solved by standard least squares procedure (4):

$$p_S^* = \arg\min_P F(p).$$

$$F(p) = \sum_{i=1}^{N} \left[U_p(\mathbf{x}_i) - f(\mathbf{x}_i)\right]^2 = \sum_{i=1}^{N} err_i^2 , \quad (4)$$

where $i = 1..N$, err_i^2 is an output error square for sample $\mathbf{x}_i \in T$.

Partial derivatives of goal function $F(p)$ are defined by expressions

$$\frac{\partial F(p)}{\partial a_{rm}} = 2\sum_{i=1}^{N}\left[\frac{\partial U_p(\mathbf{x}_i)}{\partial a_{rm}} err_i\right] = \frac{4}{q_{rm}^2}\sum_{i=1}^{N}\left[(x_{im} - a_{rm})\prod_{m=1}^{M}\mu_{rm}(x_{im}) L_r(\mathbf{x}_i) err_i\right], \quad (5)$$

$$\frac{\partial F(p)}{\partial q_{rm}} = 2\sum_{i=1}^{N}\left[\frac{\partial U_p(\mathbf{x}_i)}{\partial q_{rm}} err_i\right] = \frac{4}{q_{rm}^3}\sum_{i=1}^{N}\left[(x_{im} - a_{rm})^2\prod_{m=1}^{M}\mu_{rm}(x_{im}) L_r(\mathbf{x}_i) err_i\right], \quad (6)$$

$$\frac{\partial F(p)}{\partial k_{rm}} = 2\sum_{i=1}^{N}\left[\frac{\partial U_p(\mathbf{x}_i)}{\partial k_{rm}} err_i\right] = 2\sum_{i=1}^{N}\left[x_{im}\prod_{m=1}^{M}\mu_{rm}(x_{im}) err_i\right], \quad (7)$$