

TREINAMENTO DE REDES NEUROFUZZY VIA APRENDIZADO PARTICIPATIVO

MICHEL HELL*, PYRAMO COSTA JR.†, ROSÂNGELA BALLINI*, FERNANDO GOMIDE*

**Universidade Estadual de Campinas,
Av. Albert Einstein, 400 - Cidade Universitária "Zeferino Vaz"
Campinas, SP - Brasil, CEP 13083-970*

†*Pontifícia Universidade Católica de Minas Gerais,
Av. Dom José Gaspar, 500 - Coração Eucarístico,
Belo Horizonte, MG - Brasil, CEP 30535-610*

Emails: mbhell@dca.fee.unicamp.br, pyramo@pucminas.br, ballini@eco.unicamp.br,
gomide@dca.fee.unicamp.br

Abstract— This paper introduces a new approach to train a class of neurofuzzy networks based on the idea of participatory learning. Participatory learning is a mean to learn and revise beliefs based on what is already known or believed. The performance of the approach is verified with the well known Box and Jenkins gas furnace modeling problem, and with a short-term load forecasting problem using actual data. Comparisons with alternative training procedures suggested in the literature are included to shown the effectiveness of participatory learning to train neurofuzzy networks.

Keywords— Participatory Learning, Fuzzy Systems, Neurofuzzy Networks.

Resumo— Neste trabalho é apresentada uma nova abordagem para o treinamento de uma classe de redes neurofuzzy baseada na idéia de aprendizado participativo. O aprendizado participativo é uma forma de aprender e revisar crenças a respeito de um determinado ambiente baseado no que já se sabe a respeito deste ambiente, ou seja, uma nova observação só é incorporada ao processo de aprendizado se ela for, de certa forma, compatível com o que já se aprendeu a respeito do ambiente. A performance da abordagem é verificada através de sua aplicação a dois problemas bem conhecidos: o problema de modelagem do forno a gás de Box e Jenkins, e o problema de previsão de carga de curto prazo. Com a finalidade de mostrar a eficiência do aprendizado participativo aplicado à redes neurofuzzy são apresentados também estudos comparativos desta abordagem com métodos alternativos propostos na literatura.

Keywords— Aprendizado Participativo, Sistemas Fuzzy, Redes Neurofuzzy.

1 Introdução

Sistemas neurofuzzy vêm sendo largamente utilizados nas mais diversas áreas do conhecimento, tais como economia, matemática difusa, teoria de jogos, modelagem de sistemas, entre tantas outras.

Estes sistemas combinam os dois maiores paradigmas da inteligência computacional: redes neurais e sistemas fuzzy. Essa combinação resulta em sistemas que integram o tratamento de informação incerta provida pelos sistemas fuzzy e a habilidade de aprendizado proporcionada pelas redes neurais. A idéia principal por trás dos sistemas neurofuzzy consiste na capacidade de ajustar um sistema de inferência fuzzy a partir de um processo de treinamento baseado em dados. Assim, a definição das metodologias de treinamento a serem utilizadas está entre um dos mais importantes tópicos no desenvolvimento de sistemas neurofuzzy (Figueiredo et al., 2004).

Nos últimos anos vários métodos vem sendo propostos para o treinamento de sistemas neurofuzzy, em especial de redes neurofuzzy. Em geral estes métodos são baseados no método do gradiente descendente e/ou em estratégias de reforço associativo (Figueiredo et al., 2004; Pedrycz and Gomide, 1998; Wang, 1994; Lin and Lee, 1996).

Essas técnicas produzem resultados satis-

fatórios em muitos casos, mas, atualmente, não existe um consenso em qual deles é o mais indicado, uma vez que estes dependem das características dos dados de treinamento (Silva et al., 2005).

Em 1990 uma nova abordagem, chamada "*Participatory Learning*" (PL), foi introduzida por Yager como um modelo de aprendizado que captura muitas das características inerentes ao aprendizado humano (Yager, 1990). Essa abordagem assume que a incorporação de novas crenças a respeito de um determinado ambiente depende do que já se sabe a respeito deste ambiente (Yager, 2004).

Recentemente, o paradigma de aprendizado participativo (PL) foi utilizado para desenvolver um eficiente algoritmo de agrupamento fuzzy não supervisionado (Silva et al., 2005) e, também, para encontrar estruturas de base de regras em procedimentos de modelagem fuzzy evolutiva (Lima et al., 2006).

Este trabalho, sugere o uso da idéia de PL para desenvolver um novo procedimento de treinamento para uma classe de redes neurofuzzy híbridas. Este procedimento não depende tão fortemente da característica do conjunto de dados, uma vez que o conhecimento que já se tem a respeito do ambiente faz parte do processo de treinamento e influencia na forma com a qual novas observações

serão usadas no aprendizado (Silva et al., 2005).

Este artigo está organizado da seguinte forma: a próxima seção revisa as principais idéias a respeito do aprendizado participativo. A Seção 3 apresenta a estrutura da rede neurofuzzy utilizada. A Seção 4 detalha o processo de treinamento sugerido neste trabalho. Na Seção 5 são apresentados os resultados de simulação e os estudos comparativos. A Seção 6 conclui o trabalho, resumindo suas contribuições e propondo tópicos para investigações futuras.

2 Aprendizado Participativo

A característica principal do aprendizado participativo é que o impacto de uma observação exógena no aprendizado ou na revisão das crenças depende de sua compatibilidade com as crenças atuais. Em particular, uma observação conflitante com as crenças atuais é desconsiderada (Yager, 1990).

Seja $v \in [0, 1]^n$ a variável que representa a crença de um sistema (tudo que já se aprendeu a respeito de um determinado ambiente). O objetivo do aprendizado participativo é ajustar o valor desta variável baseado em uma seqüência de novas observações $x^k \in [0, 1]^n$ que codificam um novo conhecimento a respeito do valor da variável v . Neste sentido, x^k é uma manifestação do valor de v na k -ésima observação. Assim, usa-se o vetor x^k como um meio para “aprender” as avaliações de v . O processo de aprendizado é participativo se a contribuição de cada observação x^k para o processo de aprendizado depende de sua aceitação pelos valores correntes de v (Yager, 1990). Isto significa que, para ser relevante para o processo de aprendizagem, x^k deve estar próximo de v^k . Um mecanismo para atualização das crenças correntes pode ser dado por:

$$v^{k+1} = v^k + \alpha \rho_k (x^k - v^k) \quad (1)$$

onde $k = 1, \dots, P$, e P é o número de observações; v^{k+1} é a nova crença a respeito do sistema; v^k é a crença corrente; x^k é a observação corrente; $\alpha \in [0, 1]$ é a taxa de aprendizado; e $\rho_k \in [0, 1]$ é o grau de compatibilidade entre x^k e v^k . Uma das formas de se calcular ρ_k é dada por:

$$\rho_k = 1 - \frac{1}{n} \sum_{i=1}^n d_i^k \quad (2)$$

sendo $d_i^k = |x_i^k - v_i^k|$. Note que, se $\rho_k = 0$ (observação muito longe da crença corrente) tem-se, a partir de (1) que $v^{k+1} = v^k$ e o sistema está totalmente fechado para o aprendizado. Por outro lado, se $\rho_k = 1$ tem-se a partir de (1) que $v^{k+1} = v^k + \alpha(x^k - v^k)$, e portanto o sistema está totalmente aberto para o aprendizado. Entretanto, para que ρ_k seja igual a 1 é preciso que $x_i^k - v_i^k = 0$ para todo i . Esta condição implica que $v^k = x^k$ e assim tem-se que $v^{k+1} = v^k$, o

que implica que não ocorre nenhum aprendizado (Yager and Filev, 1993).

Um ponto que pode ser levantado a respeito do processo de treinamento descrito é que este ignora a situação em que uma seqüência de baixos valores de ρ_k 's são observados durante um longo período de tempo. Neste caso, o sistema deve se tornar mais aberto para aprender com as novas observações, uma vez que para este caso é a crença que se tem a respeito do sistema que pode estar errada e não as novas observações. Esta informação pode ser traduzida em um índice de “alerta”, usado para influenciar o processo de aprendizado, como mostrado na Figura 1. Quanto maior o índice de alerta, menor é a confiança que se tem a respeito da crença atual do sistema e observações conflitantes se tornam importantes no processo de atualização das crenças. Considere $a_k \in [0, 1]$ como

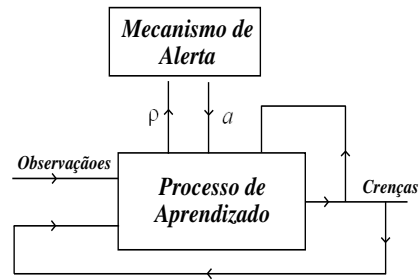


Figura 1: Aprendizado Participativo com Alerta.

sendo o índice de alerta. Quanto mais alto for o valor de a_k mais estimulado está o sistema. A dinâmica de adaptação do índice de alerta é dada por:

$$a_{k+1} = a_k + \lambda((1 - \rho_{k+1}) - a_k) \quad (3)$$

onde $\lambda \in [0, 1]$ é uma constante que controla a taxa com a qual a alerta é ajustada; quanto mais próximo λ está de 1, mais sensível fica o sistema à variações de compatibilidade. O índice de alerta pode ser visto como o complemento da confiança na crença atualmente presente no sistema.

A expressão (1) pode então ser reescrita de forma a incorporar o mecanismo de alerta da seguinte forma:

$$v^{k+1} = v^k + \alpha \rho_k^{1-a_k} (x^k - v^k) \quad (4)$$

Como pode ser verificado em (4), enquanto ρ_k mede o quanto o sistema muda sua credibilidade em suas próprias crenças, o índice de alerta a_k atua como um crítico que avisa quanto uma crença deve ser modificada frente a novas evidências.

3 Modelo da Rede Neurofuzzy

Esta seção apresenta a estrutura completa da rede neurofuzzy híbrida utilizada neste trabalho. Esta rede é composta por duas partes. A primeira parte

possui uma camada de entrada e duas camadas intermediárias e representa um sistema de inferência fuzzy. A segunda parte é formada por uma rede neural clássica de apenas uma camada que tem o propósito de agregar as saídas do sistema de inferência fuzzy, fornecendo uma saída para a rede. A Figura 2 apresenta uma rede neurofuzzy com n entradas e m saídas.

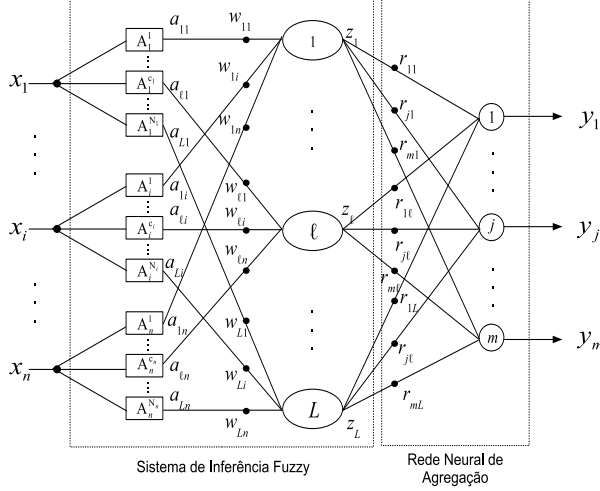


Figura 2: Estrutura da Rede Neurofuzzy Híbrida

A camada de entrada apenas fornece as entradas para os neurônios de fuzzyficação da próxima camada. A primeira camada intermediária é formada por neurônios cujas funções de ativação são as funções de pertinência dos conjuntos fuzzy que formam o particionamento do espaço de entrada. Para cada dimensão x_i^k de um vetor de entrada n -dimensional \mathbf{x}^k existem N_i conjuntos fuzzy $A_i^{\lambda_i}$, $\lambda_i = 1, \dots, N_i$ cujas funções de pertinência são as correspondentes funções de ativação dos neurônios da camada de entrada. A variável k denota o tempo discretizado, isto é, $k = 1, 2, \dots$, e será omitida no decorrer do artigo para simplificar a notação. Assim, as saídas da primeira camada intermediária são os graus de pertinência associados aos valores das entradas, isto é, $a_{\ell i} = \mu_{A_i^{\lambda_i}}(x_i)$, $i = 1, \dots, n$ e $\ell = 1, \dots, L$; onde L é o número de neurônios na segunda camada intermediária.

Os neurônios da segunda camada intermediária são neurônios lógicos chamados neurônios OU (Figura 3) cujas entradas $a_{\ell i}$ são ponderadas pelos pesos $w_{\ell i}$ onde $w_{\ell i}$ corresponde à credibilidade da entrada $a_{\ell i}$. Os neurônios lógicos são implementados utilizando as normas triangulares, isto é, o operador OU é representado por uma s -norma; diversas definições destas normas encontram-se em (Pedrycz and Gomide, 1998).

Neurônios fuzzy trabalham com graus de pertinência $a_{\ell i}$ (entradas) e pesos $w_{\ell i}$ no intervalo $[0, 1]$, produzindo, assim, um mapeamento não-

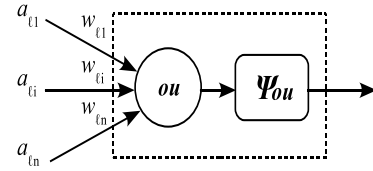


Figura 3: Neurônios Lógicos OU

linear $[0, 1]^n \rightarrow [0, 1]$. A função de ativação ψ_{OU} de um neurônio OU é, em geral, uma função não linear, mas, neste trabalho, será considerada como a função identidade, ou seja $\psi_{OU}(u) = u$.

É fácil verificar que a estrutura da rede representa um conjunto de regras *se-então* $R = \{R_k, k = 1, \dots, L\}$ da forma:

- R_k : Se (x_1 é $A_1^{\lambda_1}$ com credibilidade $w_{\ell 1}$) ...
 OU (x_i é $A_i^{\lambda_i}$ com credibilidade $w_{\ell i}$) ...
 OU (x_n é $A_n^{\lambda_n}$ com credibilidade $w_{\ell n}$)
Então, z é z_ℓ .

onde

$$z_\ell = \mathbf{S}_{i=1}^n (w_{\ell i} \mathbf{t} a_{\ell i}).$$

Desta forma é possível verificar que existe uma forte correspondência entre a estrutura da primeira parte da rede neurofuzzy e um conjunto de regras fuzzy, ou alternativamente, uma base de regras fuzzy. Além disto, o processamento do esquema induzido pela estrutura da rede está de acordo com os princípios da teoria de conjuntos fuzzy e de raciocínio aproximado (Pedrycz and Gomide, 1998).

Conforme mencionado anteriormente, a segunda parte da estrutura da rede neurofuzzy é formada por uma rede neural clássica composta por uma camada de neurônios convencionais cujas saídas y_j são uma agregação das entradas z_ℓ e os pesos $r_{j\ell}$, $\ell = 1, \dots, L$, $j = 1, \dots, m$.

O processamento da rede neurofuzzy pode então ser resumido como se segue:

1. N_i é o número de conjuntos fuzzy que constitui a partição da i -ésima entrada;
2. $a_{\ell i} = \mu_{A_i^{\lambda_i}}(x_i)$ é o grau de pertinência de x_i no conjunto fuzzy $A_i^{\lambda_i}$, sendo $a_{\ell i}$ a i -ésima entrada do neurônio ℓ da segunda camada intermediária;
3. z_ℓ é a saída do ℓ -ésimo neurônio da segunda camada intermediária;

$$z_\ell = \mathbf{S}_{i=1}^n (w_{\ell i} \mathbf{t} a_{\ell i}) \quad (5)$$

4. y_j é a saída do j -ésimo neurônio não linear

da camada de saída sendo calculada como:

$$y_j = f(u_j) = f\left(\sum_{\ell=1}^L (r_{j\ell} z_\ell)\right) \quad (6)$$

onde, $f : \mathbb{R}^L \rightarrow [0, 1]$ é uma função não linear monotonicamente crescente. Neste trabalho usou-se a função logística $f(u_j) = 1/(1 + \exp(-u_j))$;

5. $w_{\ell i}$ é o peso entre o ℓ -ésimo neurônio OU e o i -ésimo neurônio da primeira camada intermediária;
6. $r_{j\ell}$ é o peso entre a saída y_j da rede e o ℓ -ésimo neurônio da segunda camada intermediária;

A arquitetura apresentada tem como vantagens a geração automática da topologia da rede, flexibilidade quanto à utilização de diversas normas triangulares e a possibilidade de extração de regras diretamente da estrutura.

4 Processo de Treinamento

O processo de treinamento sugerido neste trabalho envolve três fases principais. A primeira fase usa o algoritmo fuzzy c-means para granularizar o espaço de entrada. A próxima fase usa o gradiente descendente para ajustar os pesos relacionados a rede neural clássica ($r_{j\ell}$'s). A última fase usa o paradigma do aprendizado participativo sumariado na seção 2 para ajustar os pesos conectados aos neurônios OU ($w_{\ell i}$'s). Os passos essenciais são detalhados a seguir.

4.1 Geração das Funções de Pertinência

Para gerar as funções de pertinência da primeira camada intermediária da rede mostrada na Figura 2 usa-se o algoritmo de agrupamento fuzzy c-means. Informações a respeito dos espaços de entrada e saída são incluídas no processo de agrupamento. O valor modal das funções de pertinência são as projeções dos centros dos grupos em seus respectivos universos. Além disso, adotou-se funções de pertinência gaussianas, onde as dispersões das gaussianas são ajustadas de forma a se obter uma distribuição cognitiva no espaço de entrada (Pedrycz and Gomide, 1998).

4.2 Ajuste dos Pesos

O primeiro passo no processo de ajuste dos pesos é estimar a saída da rede $\hat{y} \in [0, 1]^m$ para uma dada entrada $x \in [0, 1]^n$. Isto corresponde a fuzzyficar o padrão de entrada e calcular sucessivamente as saídas das camadas posteriores, usando as equações (5) e (6). Após esse passo, o processo de treinamento tem como objetivo minimizar uma medida

de erro entre a saída estimada e a saída desejada, para cada um dos padrões de entrada, isto é, minimizar

$$e = \frac{1}{2} \sum_{j=0}^m (y_j - \hat{y}_j)^2 \quad (7)$$

onde \hat{y}_j é o valor da j -ésima componente do vetor de saída estimada e y_j é a j -ésima componente da saída desejada.

A atualização dos pesos na camada de saída é feita utilizando o gradiente descendente, isto é

$$\Delta r_{j\ell} = \eta (y_j - \hat{y}_j) f'(u_j) z_\ell \quad (8)$$

onde $f'(u_j) = f(u_j)(1 - f(u_j))$ é a derivada da função de ativação avaliada no ponto u , $u_j = \sum_{\ell=1}^L (r_{j\ell} z_\ell)$, e η é a taxa de aprendizado.

O próximo passo é o ajuste dos pesos dos neurônios lógicos usando a idéia do aprendizado participativo, conforme apresentado a seguir.

O Procedimento do Aprendizado Participativo

Como discutido na Seção 3, o sistema de inferência fuzzy representado pela estrutura de neurônios lógicos pode ser visto como uma relação fuzzy da forma $G : A \times Z \rightarrow [0, 1]$ onde A e Z são dois universos de discursos finitos e G é um subconjunto do produto cartesiano destes dois universos. Em particular, tem-se que para $a^k \in A$ e $G \subset A \times Z$ é possível calcular um elemento $z^k \in Z$ relacionado com a^k da forma $z^k = a^k \circ G$.

Se a s -norma dos neurônios OU forem escolhidas como sendo a função máximo, tem-se que o operador 'o' torna-se max_t (sup_t) e para um dado par (a^k, z^k) a relação G pode ser encontrado resolvendo-se o problema de estimação relacional fuzzy, cuja solução é dada por (Pedrycz and Gomide, 1998):

$$G = a^k \overset{T}{\varphi} z^k \quad (9)$$

onde $\overset{T}{\varphi}$ denota a transposta e φ é tal que:

$$(a_{\ell i}^k \varphi z_\ell^k) = sup(c \in [0, 1] \mid a_{\ell i}^k t c \leq z_\ell^k) \quad (10)$$

É possível verificar que a matriz relacional $G = [g_{\ell i}]$ tem uma forte relação com a matriz de pesos $W = [w_{\ell i}]$. De fato, para um dado padrão $[x_k, y_k]$ e uma matriz $R = [r_{j\ell}]$ de pesos associados a camada de saída, a melhor aproximação para o mapeamento $f : x_k \rightarrow y_k$ obtido pela rede neurofuzzy ocorre quando $W \equiv G$.

Para encontrar a relação G é necessário inicialmente se obter os valores de saída dos neurônios lógicos z_k . Estes valores podem ser estimados a partir de y_k e R resolvendo-se o problema de mínimos quadráticos linear restrito:

$$\min_{z^k} \left\{ \frac{1}{2} |Rz^k - f^{-1}(y^k)|^2 \right\} \text{ sujeito a } 0 \leq z^k \leq 1 \quad (11)$$

onde $f^{-1}(\cdot)$ é a inversa da função f em (6).

Assim, é possível calcular a relação fuzzy G usando (10).

Uma vez que G tenha sido calculada, pode-se ajustar os pesos $w_{\ell i}$'s baseado no paradigma do aprendizado participativo, da seguinte forma.

$$\Delta W^k = \alpha (\rho^k)^{1-a_k} (G^k - W^k) \quad (12)$$

onde α é a taxa de aprendizado; $k = 1, \dots, P$ e P é o número de padrões de treinamento; a_k é calculado de acordo com (3), e

$$\rho^k = \frac{1}{Ln} \sum_{\ell=1}^L \sum_{i=1}^n |g_{\ell i}^k - w_{\ell i}^k|. \quad (13)$$

O processo de treinamento da rede neurofuzzy usando a abordagem PL (NFPL) proposta neste trabalho pode ser resumida da seguinte forma:

INÍCIO NFPL

```

Ler os dados de Treinamento;
Granularizar o Espaço de Entrada (seção 4.1);
p:= n° de padrões de treinamento;
ENQUANTO err > tolerância do erro FAÇA
  t:=1;
  ENQUANTO t ≤ p FAÇA
    Leia ( $x^k, y^k$ );
    Calcule a saída de rede  $\hat{y}^k$  (seção 3);
    Calcule o erro de aproximação usando (7);
    Calcule a variação dos pesos  $r_{ji}$ 's usando (8);
    Atualize os pesos  $r_{ji}$ 's;
    t:=t+1;
  FIM ENQUANTO
  t:=1;
  ENQUANTO t ≤ p FAÇA
    Leia ( $x^k, y^k$ );
    Calcule  $z^k$  a partir de  $\hat{y}^k$  e R usando (11);
    Calcule a relação fuzzy G usando (9);
    Determine o grau de compatibilidade  $\rho_k$  usando (13);
    Determine o índice de estimulação usando (3);
    Calcule a variação dos pesos  $w_{\ell i}$  usando (12);
    Atualize os pesos  $w_{\ell i}$ 's;
    t:=t+1;
  FIM ENQUANTO
  Calcule o erro geral de aproximação err
FIM ENQUANTO
FIM NFPL

```

5 Resultados Experimentais

Para se avaliar o desempenho do algoritmo de treinamento introduzido neste trabalho, este foi comparado com abordagens alternativas sugeridas na literatura. Os exemplos apresentados nesta seção incluem identificação de sistemas e predição de séries temporais. Para ambos os exemplos adotou-se como s -norma a função \max e o produto algébrico como t -norma para os neurônios OU apresentados na Figura 3.

O Forno a Gás de Box e Jenkins

O problema de identificação do forno a gás proposto por Box e Jenkins é um bem conhecido

problema de referência. O processo de identificação usa 290 pares entrada/saída obtidos a partir de um forno a gás de laboratório (Box and Jenkins, 1976). Cada amostra consiste das taxas de fluxo de metano (entradas x^k do sistema) e a concentração de CO_2 na saída do forno (saídas y^k do sistema). Este é um processo dinâmico com uma entrada x^k e uma saída y^k . O objetivo é prever a saída corrente y^k a partir de entradas e saídas em instantes anteriores, com o menor erro possível. Diferentes estudos indicam que a melhor estrutura do modelo para este sistema é:

$$y^k = f(y^{k-1}, x^{k-4}) \quad (14)$$

A Figura 4 apresenta o resultado obtido pela abordagem NFPL para o problema proposto. Resultados comparativos com outras abordagens fuzzy/neurofuzzy utilizadas para resolver o mesmo problema a partir das mesmas entradas são apresentados na Tabela 1.

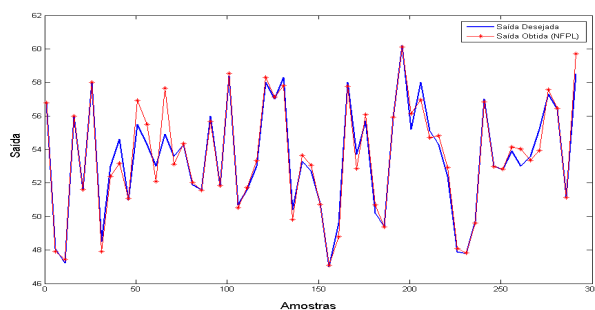


Figura 4: Identificação do Forno a Gás de Box e Jenkins

Tabela 1: Identificação do Forno a Gás

Modelo	No de Regras	Erro Médio Quadrático
Tong (Tong, 1980)	19	0.6848
Pedrycz (Pedrycz, 1984)	81	0.5656
Xu an Lu (Xu and Lu, 1987)	25	0.5727
Delgado (Delgado et al., 1997)	4	0.4100
Yoshinari (Yoshinari et al., 1993)	6	0.5460
NFPL	9	0.4528

Previsão de Carga

Um exemplo que mostra a eficiência do aprendizado participativo no treinamento neurofuzzy é o problema de previsão de carga de curto prazo. O objetivo da previsão de carga de curto prazo é prever as 24 horas do dia seguinte.

Para o treinamento foi utilizado um conjunto de dados composto por dados de carga coletados de hora em hora no período de maio de 2000 a fevereiro de 2001 de um sistema elétrico de potência localizado na região Sudeste do Brasil. As variáveis de entrada do modelo são L_{h-1} e L_{h-2} , que

representam a carga nos instantes $h - 1$ e $h - 2$, respectivamente. A Figura 5 mostra o resultado da previsão feita utilizando-se a abordagem NFPL para o dia 12 de setembro de 2000.

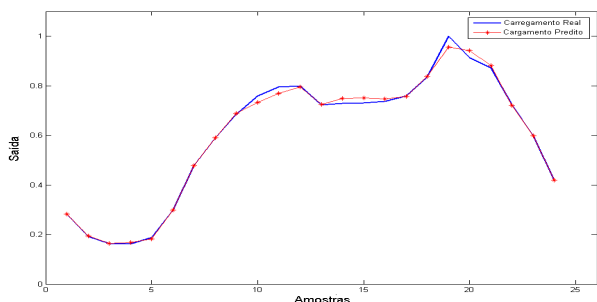


Figura 5: Previsão de Carga de Curto Prazo

Nesta aplicação foram definidos 8 conjuntos fuzzy para cada uma das variáveis de entrada, o que gera um total de 64 neurônios na segunda camada intermediária da rede (64 regras fuzzy).

Com a finalidade de comparar os resultados obtidos pela rede NFPL com outras abordagens, mais dois modelos foram aplicados ao mesmo problema: uma rede neural de múltiplas camadas clássica e um sistema neurofuzzy adaptativo (Jang, 1993) com o mesmo número de conjuntos fuzzy para granularizar cada entrada. Para avaliar os resultados, foi usado o erro absoluto médio dado por:

$$MAE = \frac{1}{P} \sum_{k=1}^P \frac{|\hat{y}^k - y^k|}{y^k} \quad (\%) \quad (15)$$

onde \hat{y}^k é o k -ésimo valor predito, y^k é k -ésimo valor desejado e P é o número de valores a prever. Os resultados obtidos são apresentados na Tabela 2.

Tabela 2: Previsão de Carga

Modelo	MAE %
Perceptron de Múltiplas Camadas (MLP)	7.20
Sistema Neurofuzzy Adaptativo (ANFIS)	5.74
NFPL	3.98

6 Conclusões

Neste trabalho foi introduzido um novo procedimento para o treinamento de uma classe de redes neurofuzzy híbridas baseadas no paradigma do aprendizado participativo. Resultados experimentais mostraram que o aprendizado participativo fornece uma forma atrativa para o treinamento supervisionado de redes neurofuzzy que não depende tão fortemente das características do conjunto de dados de treinamento e ainda fornece uma alternativa eficiente para as abordagens fuzzy/neurofuzzy propostas na literatura.

Além disso, os resultados obtidos indicam que em estudos futuros esta nova abordagem pode

facilmente ser estendida para outras classes de rede neurofuzzy, incluindo redes com recorrências locais e globais.

Agradecimentos

Os autores agradecem à FAPESP, a Fundação de Amparo a Pesquisa do Estado de São Paulo, pelos apoios 03/05042-1 e 03/10019-9, e ao CNPq, o Conselho Nacional de Pesquisa e Desenvolvimento, pelo apoio 304857/2006-8.

Referências

- Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis- Forecasting and Control*, second edn, Holden Day, CA, USA.
- Figueiredo, M., Ballini, R., Soares, S., Andrade, M. and Gomide, F. (2004). Learning algorithms for a class of neurofuzzy network and application, *IEEE Trans. on System, Man and Cybernetics - Part C* **34**(3): 293–301.
- Jang, R. (1993). Anfis: Adaptive network based fuzzy inference system, *IEEE Trans. on Systems Man and Cybernetics* **23**(3).
- Lima, E., Gomide, F. and Ballini, R. (2006). Participatory evolving fuzzy modeling, *Proc. 2006 International Symposium on Evolving Fuzzy Systems*, Lake district, UK, pp. 36–41.
- Lin, C. T. and Lee, C. S. (1996). *Neural Fuzzy Systems*, Prentice-Hall, NJ, USA.
- Pedrycz, W. and Gomide, F. (1998). *An Introduction of Fuzzy Sets: Analysis and Design*, MIT Press, Cambridge, MA, USA.
- Silva, L., Gomide, F. and Yager, R. R. (2005). Participatory learning in fuzzy clustering, *Proc. 14th IEEE International Conference on Fuzzy Systems*, Reno, USA, pp. 857–861.
- Wang, L. X. (1994). *Adaptive Fuzzy Systems and Control*, Prentice-Hall, NJ, USA.
- Yager, R. R. (1990). A model of participatory learning, *IEEE Trans. on System, Man and Cybernetics* **20**(5): 1229–1234.
- Yager, R. R. (2004). Participatory learning: A paradigm for more human like learning, *Proc. 2004 IEEE International Conference on Fuzzy Systems*, Vol. 1, pp. 79–84.
- Yager, R. R. and Filev, D. P. (1993). A fuzzy logic controller view of participatory learning, *Proc. Second IEEE International Conference on Fuzzy Systems*, Vol. 2, pp. 912–917.