

Classificação de Padrões em Séries Temporais Utilizando Um Sistema Neuro-Difuso

Toni Roberto Gomes Pimentel¹, Sandra Aparecida Sandri², Fernando Manuel Ramos²

¹Programa de Mestrado em Computação Aplicada – CAP
Instituto Nacional de Pesquisas Espaciais – INPE

²Laboratório Associado de Computação e Matemática Aplicada – LAC
Instituto Nacional de Pesquisas Espaciais – INPE

tonipimentel@gmail.com.br

{sandri, fernando}@lac.inpe.br

***Abstract.** Although, in general, the studies on time series are divided into clustering and future value prediction, we focused our studies on time series subsequences classification into pre-defined classes, based on previous information about the evolutionary behaviour of specific scenarios. We propose the use of a neuro-fuzzy inference model to this classification. Typically, these models approximate non-linear functions through a combination of linear local models using approximate information and uncertainty.*

***Resumo.** Apesar de, em geral, os estudos em séries temporais serem divididos em agrupamento de padrões e determinação de valor futuro, focamos nosso estudo em classificação de subsequências de uma série temporal em classes pré-definidas, com base em informações a priori sobre o comportamento evolutivo das séries para cenários específicos. Nós propomos a utilização de um modelo de inferência neuro-difuso para essa classificação. Tipicamente, estes modelos aproximam funções não lineares através da combinação de modelos lineares locais usando de incerteza e informações aproximadas.*

Palavras-chave: *Séries Temporais, Classificação de Padrões, Sistemas Neuro-Difusos*

1. Introdução

Estatística material no mundo real muitas vezes toma a forma de uma sequência de dados, indexados pelo tempo. Esses dados são referidos como séries temporais e ocorrem em diversas áreas da atividade humana: preços de ações nos mercados financeiros, fenômenos astrofísicos como manchas solares, números de vendas de uma empresa, informações demográficas de uma entidade geográfica, entre outros. Medições em séries temporais podem ser contínuas – feitas continuamente no tempo – ou discretas, isto é, em intervalos específicos, geralmente igualmente espaçados. Duas vertentes aparecem em análise de séries temporais: *clustering*, ou agrupamento de padrões, e estimação de valor futuro. A essência da análise de séries temporais é que existem padrões de comportamento repetidos que podem ser identificados e modelados. A repetição, seja de comportamento suave ou turbulento, é essencial para a generalização [Popoola et al. 2007].

Agrupamento e estimação de valor futuro são as áreas de estudo mais abrangidas em séries temporais. Entretanto, este trabalho não pretende investir em nenhuma dessas áreas, mas

na classificação de padrões. A diferença entre classificação e agrupamento é clara, mas não evidente: na classificação os dados devem ser atribuídos a grupos pré-definidos, já no processo de agrupamento os grupos, ou *clusters*, são o resultado da análise dos dados [Oliveira 2011]. Sendo assim, a grande diferença entre esses dois estudos é a presença de informação a priori sobre os padrões das séries. Ou seja, existe conhecimento especialista sobre como a série evolui para determinados cenários. Fazendo um contraponto, podemos determinar qual o cenário existente para certo subconjunto de observações de acordo com o comportamento dos dados desse subconjunto.

Modelos de aprendizagem supervisionada são os mais indicados para essa tarefa, já que seu algoritmo analisa os dados de treinamento e produz uma função, que é chamada de classificador. O uso de inteligência computacional torna-se ideal na abordagem do problema, já que esses métodos são uma promessa de processamento eficiente e de comportamento tolerante a falhas [Popoola et al. 2007].

O objetivo principal deste trabalho é classificar padrões em séries temporais utilizando uma técnica que combina redes neurais artificiais e lógica difusa, fusão mais conhecida como sistema neuro-difuso. Essa combinação resulta em um sistema híbrido inteligente que combina o estilo de raciocínio humano de sistemas difusos com a estrutura conexionista e de aprendizagem de redes neurais.

2. Séries Temporais

Uma série temporal pode ser representada como um conjunto de observações $X_t : t = 1, \dots, N$, onde t é o índice de tempo e N é o número total de observações [Morettin and De Castro Toloí 1981]. Em geral, as observações de séries temporais são autocorrelacionadas. Essa correlação resulta em padrões que são úteis para a análise dos dados. As séries temporais são divididas em componentes, e são funções deles:

$$X_t = f(T_t, S_t, C_t, I_t) \quad (1)$$

onde T_t, S_t, C_t e I_t representam, respectivamente, os componentes, no tempo t : tendência, sazonal, cíclico e aleatório [Souza 1989].

O componente tendência representa o incremento, ou decréscimo, gradual nos valores das observações de uma série temporal. O componente sazonal representa as flutuações da série de acordo com algum fator de sazonalidade. O componente cíclico apresenta um comportamento similar ao componente sazonal, embora tenha normalmente comprimento maior que este, pelo fato de não apresentar duração uniforme. Por fim, o componente aleatório é usado para admitir comportamento não determinístico no modelo [Makridakis and Wheelwright 1989].

A análise de séries temporais se concentra em três objetivos básicos: previsão de progressões de curto prazo, modelagem de comportamento a longo prazo e caracterização de propriedades subjacentes [Gershenfeld and Weigend 1993]. Modelos convencionais, para testar a hipótese de que relações complexas e potencialmente causais existem entre vários elementos de uma série temporal, baseados em métodos paramétricos, expressam estas relações através de uma variedade de maneiras. O mais popular é o modelo autorregressivo onde existe uma hipótese de que a causalidade conecta o valor da série no tempo t para os seus p valores anteriores. Por outro lado, as técnicas de inteligência

computacional, como sistemas difusos, algoritmos genéticos, redes neurais artificiais e sistemas híbridos, não fazem suposições sobre a estrutura dos dados, mas sim, fazem uma “aproximação universal” para fornecer o mapeamento não linear de funções complexas [Kosko 1994], [Wang and Mendel 1992] e [Ying 1998].

3. Sistemas de Inferência Difusos em Séries Temporais

As séries temporais são, invariavelmente, não estacionárias, e as suposições sobre suas estruturas feitas por meio de métodos tradicionais são difíceis de verificar [Moorthy et al. 1998], tornando esses métodos inadequados para sistemas mesmo que “moderadamente complicados” [Gershenfeld and Weigend 1993]. Além disso, dados do mundo real podem apresentar uma sobreposição de muitos processos, exibindo dinâmicas diversas. Cada vez mais, as técnicas de inteligência computacional, como sistemas difusos, têm sido usadas para modelar relações complexas nessas séries. Ao contrário dos métodos convencionais, os modelos de inteligência computacional são “não paramétricos” e aprendem sem fazer suposições sobre o processo de geração de dados [Berardi and Zhang 2003].

Em particular, sistemas difusos são utilizados devido a interpretabilidade das regras linguísticas geradas por esses métodos. Sistemas de inferência difusos (FIS, do inglês *Fuzzy Inference Systems*) são descritos como aproximadores universais que podem ser utilizados para modelar as relações não lineares entre entradas e saídas. A operação de um FIS normalmente depende da execução de quatro tarefas principais: a codificação – também muito chamada de fuzzificação –, inferência, composição e decodificação – também chamada de defuzzificação [Takagi and Sugeno 1985].

Existem também os modelos neuro-difusos, que incorporam pontos fortes de redes neurais, tais como aprendizagem e capacidade de generalização, e de sistemas difusos, como raciocínio qualitativo e capacidade de modelagem da incerteza. O sistema de inferência adaptativo neuro-difuso (ANFIS, do inglês *Adaptative Neuro-Fuzzy Inference System*), proposto por Jang [Jang 1993], é um dos mais utilizados métodos neuro-difusos. O ANFIS é uma rede neural que modela o tipo de sistemas de inferência difusos TSK e compreende cinco camadas, cada camada sendo funcionalmente equivalente a um FIS.

Embora o pré-processamento de séries temporais possa ser benéfica, ela também pode introduzir artefatos que não têm nem tendências, nem sazonalidade. Dada a forte natureza da dependência de dados de sistemas difusos, a introdução de artefatos através de dados de pré-processamento pode resultar em um desempenho de previsão significativamente degradado (ver Figura 1).

4. Sistemas Difusos

A teoria dos conjuntos difusos, formalizada pelo professor Lofti Zadeh na Universidade da Califórnia em 1965, foi desenvolvida para tratar do aspecto vago da informação. Ela constitui o ponto de partida no desenvolvimento de sistemas difusos. As bases desta teoria são estabelecidas no trabalho de Zadeh publicado em 1965 na revista *Information and Control* [Zadeh 1965]. Nesta publicação se introduziu, pela primeira vez de maneira formal, a definição de um conjunto difuso. Esta dá origem a uma série de conceitos, operações e medidas que são aplicáveis a inúmeras disciplinas da ciência.

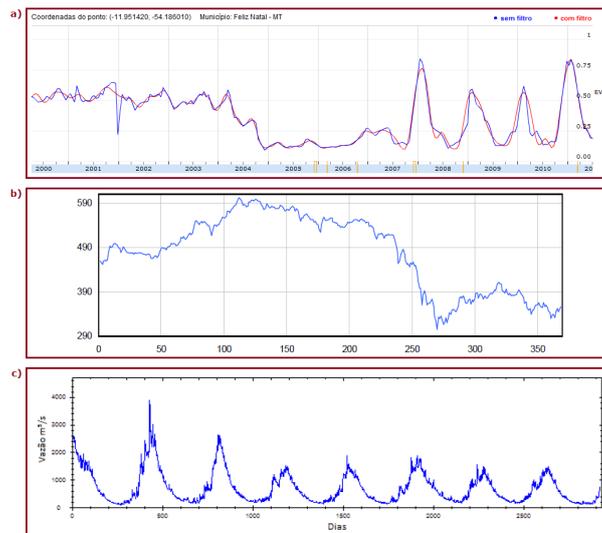


Figura 1. Exemplos de séries temporais: a) Série EVI2 para coordenada selecionada (-11.951420, -54.186010). Linha azul (sem filtro *wavelet*) e vermelha (com filtro *wavelet*) Fonte: <http://www.dsr.inpe.br/laf/series/>. b) Preços de ações da IBM, de 17 de maio de 1961 a 02 novembro de 1962 [Popoola et al. 2007]. c) Gráfico das vazões de água do Rio São Francisco entre 1978 e 1986 [Batista 2010].

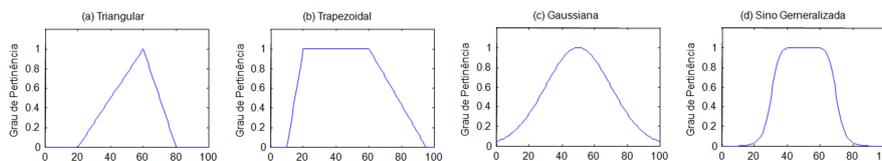


Figura 2. Funções de pertinência.

Segundo [Zadeh 1965] um conjunto difuso é definido como: “Uma classe de objetos com graus de pertinência contínuos. Tal conjunto é caracterizado por uma função de pertinência que varia entre 0 e 1”. De acordo com esta definição, um conjunto difuso A está composto por duas partes essenciais: seus elementos x pertencentes ao universo de discurso X e a função de pertinência μ_A que assina o grau de pertinência destes elementos a seu conjunto. Em outras palavras temos que um conjunto difuso A é definido matematicamente como $\mu_A : X \rightarrow [0, 1]$. Assim, $\mu_A(x) = 0$ indica x não compatível com A , $0 < \mu_A(x) < 1$ indica x parcialmente compatível com A e $\mu_A(x) = 1$ indica x totalmente compatível com A .

O tipo de representação da função de pertinência depende do conjunto base. Se este conjunto é composto de muitos valores então uma representação paramétrica é adequada. Assim, para que as funções sejam usadas, necessita-se, apenas, que parâmetros sejam adaptados. Normalmente funções de pertinência lineares são preferidas, devido à sua simplicidade e eficiência com respeito a computabilidade, principalmente funções trapezoidais e triangulares, que são definidas por quatro e três parâmetros, respectivamente. Entretanto, para algumas aplicações a modelagem requer curvas continuamente diferenciáveis e transições suaves, que os trapézios não têm (ver Figura 2).

O sistema de inferência difuso proposto por Takagi e Sugeno em [Takagi and Sugeno 1985] é uma poderosa ferramenta para a modelagem de sis-

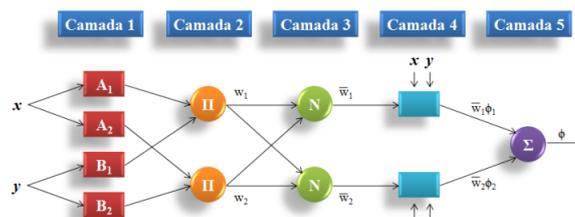


Figura 3. Exemplo de Sistema ANFIS.

temas complexos não-lineares. A modelagem Takagi-Sugeno-Kang (TSK) é uma abordagem multimodal em que submodelos lineares associados com as regras de TSK são combinados para descrever o comportamento global do sistema. Regras TSK têm elevados graus de liberdade para melhorar o seu desempenho que, conseqüentemente, torna possível expressar os comportamentos complicados com um pequeno número de regras.

Baseado na modelagem TSK e com o objetivo criar um sistema de inferência cujos parâmetros das funções de pertinência são ajustados automaticamente, surgiu, proposto por Jang [Jang 1993], o sistema de inferência adaptativa neuro-difuso (ANFIS). Utilizando um conjunto de dados entrada-saída ele constrói um sistema de inferência cujos parâmetros da função de pertinência são ajustados por meio de um algoritmo de retropropagação padrão, ou combinado a algum método como o dos mínimos quadrados. Isso permite que seus sistemas difusos aprendam com os dados que estão sendo modelados.

O modelo apresentado na Figura 3 é formado por cinco camadas, as quais serão descritas a seguir:

- Camada 1 - Sendo x e y os valores de entrada para o sistema, os nós desta camada (A_1 , A_2 , B_1 e B_2) representam o grau de pertinência destes nós em relação aos valores de entrada, ou seja, esta camada realiza a fase de codificação do sistema.
- Camada 2 - As saídas dos nós desta camada representam a “força de disparo de uma regra”. Qualquer operador de norma T que represente o operador AND lógico pode ser utilizado como função de ativação nesta camada [Jang 1993].
- Camada 3 - Cada nó desta camada funciona como um operador de normalização do nível de disparo das regras.
- Camada 4 - Esta camada calcula o produto entre as saídas da camada 3 (níveis de disparo normalizados) e os valores de entrada x e y , ponderados, os quais [Jang 1993] chamou de parâmetros consequentes.
- Camada 5 - Esta camada realiza a fase de decodificação do sistema. É formada por apenas um nó que realiza o somatório das saídas da camada 4.

5. Metodologia

Como dito anteriormente, o objetivo deste trabalho é classificar padrões em séries temporais utilizando uma técnica que combina, em um sistema, redes neurais artificiais e lógica difusa. Esses sistemas incorporam o estilo de raciocínio humano através da utilização de conjuntos difusos e um modelo linguístico que consiste de um conjunto de regras SE-ENTÃO. A força principal dos sistemas neuro-difusos é que eles são aproximadores universais com a capacidade de gerar regras SE-ENTÃO interpretáveis [Jin 2000].

A rigor, qualquer sistema que misture os dois paradigmas poderia ser chamado de neuro-difuso. No entanto, o termo é utilizado para um tipo específico de sistema que, de certa forma, engloba os dois paradigmas. Nestes sistemas, os termos e regras de um sistema difuso são aprendidos mediante a apresentação de pares (padrão de entrada - saída desejada). Eles apresentam dois comportamentos distintos, dependendo de estar numa fase de aprendizado ou numa fase de processamento da informação: na fase de aprendizado, eles têm um comportamento de redes neurais, e na fase de processamento, eles se comportam como um sistema difuso. Estes sistemas são capazes de solucionar problemas apresentados pelos paradigmas em que se baseiam [Sandri and Correa 1999].

Para a tarefa proposta utilizaremos o sistema de inferência adaptativa neuro-difuso do Matlab, o ANFIS. Ele é a implementação de um sistema neuro-difuso apresentado por Jang [Jang 1993]. Para a modelagem das regras não é necessário ter uma estrutura pré-determinada, nem saber como as funções de pertinência serão, já que o ANFIS, no período de aprendizagem, realiza a tarefa de identificar os parâmetros das funções, tanto no antecedente quanto no consequente de cada regra.

6. Aplicação em Séries Temporais de Índices de Vegetação

A utilização de produtos e técnicas de sensoriamento remoto e geoprocessamento nas análises ambientais têm se tornado uma prática cada vez mais freqüente entre as diversas áreas de pesquisa. No caso do uso da terra e da cobertura vegetal, estas técnicas contribuem de modo expressivo para a rapidez, eficiência e confiabilidade nas análises que envolvem os processos de degradação da vegetação natural, fiscalização dos recursos florestais, desenvolvimento de políticas conservacionistas, bem como vários outros fatores que podem ocasionar modificações na vegetação.

[de Freitas 2012] discute e desenvolve o conceito de laboratório virtual de sensoriamento remoto que permite aos usuários terem acesso à visualização e análise de séries temporais de dados derivados de sensores remotos. Num estudo de caso, as técnicas estudadas por ele foram aplicadas nas séries temporais de índices de vegetação EVI2 derivados de produtos de 16 dias do sensor MODIS (*Moderate Resolution Imaging Spectroradiometer*), a bordo das plataformas espaciais TERRA e AQUA. A nova abordagem de visualização de séries temporais integradas em um ambiente virtual permite a inferência sobre a dinâmica do uso e cobertura da terra.

As séries temporais de índice de vegetação disponíveis por meio do trabalho de Freitas [de Freitas 2012] possuem alguns padrões de comportamento, com cada padrão pertencendo a uma classe de uso e cobertura da terra. No exemplo da Figura 4b, pode-se observar quatro desses padrões – floresta, desmatamento, pouca vegetação e agricultura –, entretanto, muitos outros padrões podem ser observados ao longo da área abordada por esse trabalho.

Como testes preliminares do conceito de classificação de padrões em séries temporais utilizando um sistema neuro-difuso, o ANFIS, do Matlab, foi utilizado para reconhecer 6 classes de uso e cobertura da terra. Ele foi treinado com 48 padrões divididos nas classes cana-de-açúcar (Figura 5a), floresta (Figura 5b), cerrado (Figura 5c), agricultura-x (Figura 5d), agricultura-y (Figura 5e) e pastagem (Figura 5f), e validado com 24 padrões, sendo que todos esses padrões foram obtidos em 10 pontos geográficos diferentes da América do Sul.

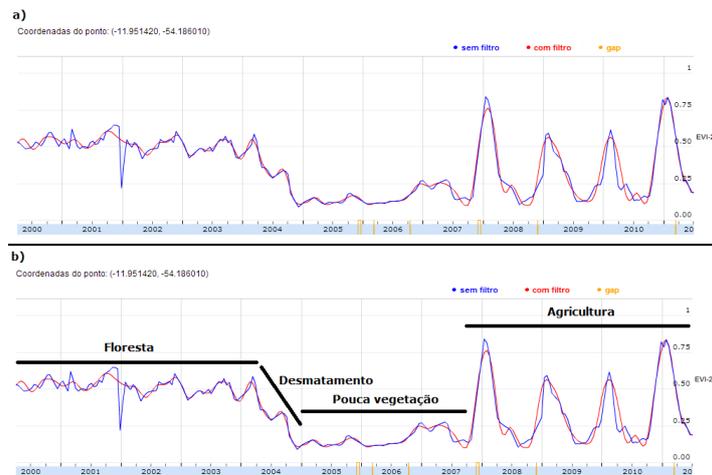


Figura 4. a) Série temporal de índices de vegetação. b) Classes presentes na série.

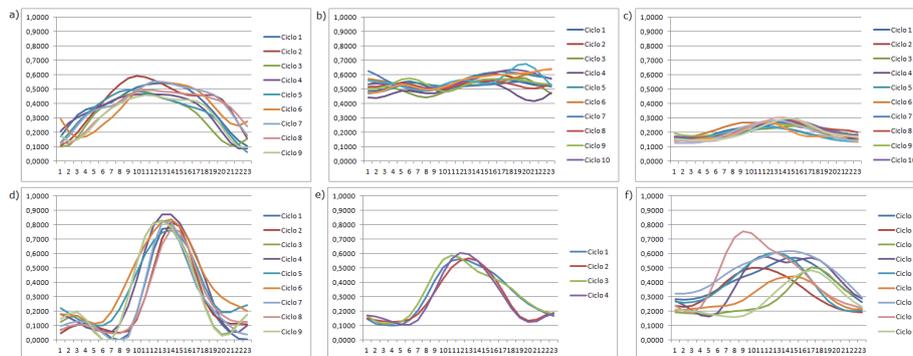


Figura 5. Classes: a) Cana-de-açúcar; b) Floresta; c) Cerrado; d) Agricultura-x; e) Agricultura-y; f) Pastagem.

O ANFIS foi capaz de, após o período de treinamento, classificar corretamente 78,95% dos novos casos (padrões) apresentados à rede, obtendo resultado promissor para o andamento da pesquisa proposta.

7. Conclusões

Sabemos que os métodos estatísticos convencionais, métodos de inteligência computacional, e também métodos híbridos, têm sido usados para caracterizar padrões de repetição de dados em séries temporais. Em particular, modelos utilizando lógica difusa representam séries temporais em termos de regras difusas. Tais métodos, baseados em regras, são considerados vantajosos porque proporcionam uma visão não só no processo de raciocínio utilizado para gerar os resultados, mas também a interpretação dos resultados obtidos. Regras difusas fornecem uma estrutura poderosa para a mineração de dados e explicação do comportamento deles, já que sistemas difusos permitem a modelagem qualitativa, com o uso de incerteza e informações aproximadas.

Técnicas neuro-adaptativas de aprendizagem proporcionam um método para o procedimento de modelagem difusa para aprender informações sobre um conjunto de dados. Uma vantagem do ANFIS é que ele calcula os parâmetros da função de pertinência que melhor permitem o sistema de inferência difusa rastrear os dados de entrada-saída.

Referências

- Batista, R. (2010). Busca de padrões em séries temporais.
- Berardi, V. and Zhang, G. (2003). An empirical investigation of bias and variance in time series forecasting: modeling considerations and error evaluation. *Neural Networks, IEEE Transactions on*, 14(3):668–679.
- de Freitas, R. M. (2012). Laboratório virtual para visualização e caracterização do uso e cobertura da terra utilizando imagens de sensoriamento remoto.
- Gershenfeld, N. and Weigend, A. (1993). The future of time series: Learning and understanding. *Working Papers*.
- Jang, J. (1993). Anfis: Adaptive-network-based fuzzy inference system. *Systems, Man and Cybernetics, IEEE Transactions on*, 23(3):665–685.
- Jin, Y. (2000). Fuzzy modeling of high-dimensional systems: complexity reduction and interpretability improvement. *Fuzzy Systems, IEEE Transactions on*, 8(2):212–221.
- Kosko, B. (1994). Fuzzy systems as universal approximators. *Computers, IEEE Transactions on*, 43(11):1329–1333.
- Makridakis, S. and Wheelwright, S. (1989). *Forecasting methods for management*. Wiley New York.
- Moorthy, M., Cellier, F., and LaFrance, J. (1998). Predicting us food demand in the 20th century: A new look at system dynamics. In *Proceedings of SPIE*, volume 3369, page 343.
- Morettin, P. and De Castro Tolo, C. (1981). *Modelos para previsão de séries temporais*. Instituto de matemática pura e aplicada.
- Oliveira, T. (2011). Clusterização de dados utilizando técnicas de redes complexas e computação bioinspirada.
- Popoola, A., Ahmad, S., and Ahmad, K. (2007). Fuzzy-wavelet method for time series analysis.
- Sandri, S. and Correa, C. (1999). Lógica nebulosa. *Escola de redes neurais: Conselho Nacional de Redes Neurais*, 5:73–90.
- Souza, R. (1989). *Modelos estruturais para previsão de séries temporais: abordagens clássica e bayesiana*. IMPA.
- Takagi, T. and Sugeno, M. (1985). Fuzzy identification of system and its applications to modelling and control. *IEEE Trans. Syst., Man, and Cyber*, 1:5.
- Wang, L. and Mendel, J. (1992). Generating fuzzy rules by learning from examples. *Systems, Man and Cybernetics, IEEE Transactions on*, 22(6):1414–1427.
- Ying, H. (1998). General siso takagi-sugeno fuzzy systems with linear rule consequent are universal approximators. *Fuzzy Systems, IEEE Transactions on*, 6(4):582–587.
- Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, 8(3):338–353.