

**USO DE REDES NEURAS ARTIFICIAIS E TEORIA DE CONJUNTOS APROXIMATIVOS NO ESTUDO DE PADRÕES CLIMÁTICOS SAZONAIS**

JULIANA A. ANOCHI, JOSÉ DEMISIO S. D. SILVA

*Laboratório Associado de Computação e Matemática Aplicada (LAC), Instituto Nacional de Pesquisas Espaciais (INPE)*

*Av. dos Astronautas, 1758, São José dos Campos-SP, CEP:12227-010, Brasil*

*E-mails: juliana.anochi@lac.inpe.br, demisio@lac.inpe.br*

**Abstract**—This work presents an Artificial Intelligence based approach for attribute reduction of reanalysis climate data to build a climate forecasting model using artificial neural networks. The methodology uses Rough Sets Theory for retrieving relevant information from the available data, thus reducing the correlation redundancy among the variables used for forecasting purposes. Neural network based forecasting models are developed for Northeast Brazil, by learning the seasonal behavior of the precipitation variable.

**Keywords**— Climate Forecasting, Rough Sets Theory, Artificial Neural Networks

**Resumo**— Este trabalho utiliza técnicas de Inteligência artificial baseado em um método de redução de atributos, para realização de previsão climática usando um modelo de Rede Neural Artificial. Para o desenvolvimento desta metodologia utilizou-se a Teoria dos Conjuntos Aproximativos para extrair informações relevantes dos dados, visando reduzir a redundância entre as variáveis. O processo de previsão climática foi desenvolvido sobre a região Nordeste do Brasil, para aprender o comportamento sazonal da variável de precipitação.

**Palavras-chave**— Previsão Climática, Teoria dos Conjuntos Aproximativos, Redes Neurais Artificiais.

## 1 Introdução

Este trabalho apresenta um método de redução de atributos que utiliza a Teoria de Conjuntos Aproximativos (TCA) usada em processos de extração da informação e tomada de decisão em aplicações de Inteligência Artificial (IA). O objetivo do trabalho é derivar conjuntos de dados de dimensão reduzida (redutos ou reduções) para o treinamento de redes neurais artificiais para realizar a previsão climática, a partir de um conjunto de dados climáticos sazonais de reanálise.

O método apresentado é uma proposta de abordagem do problema relacionado com grandes bases de dados, que têm tornado-se frequentes em função do aumento no número e na disponibilidade de sensores que permitem um melhor monitoramento de fenômenos físicos, como no caso do estado da atmosfera em meteorologia.

Assim, na meteorologia a previsão de clima ou tempo pode envolver um grande número de variáveis, que correspondem às medidas ou observações dos fenômenos físicos que se relacionam com os processos da atmosfera.

Quanto maior o número de informações obtidas sobre os fenômenos físicos atmosféricos, mais subsídios existirão para tentar melhorar a previsão. Mas isso implica diretamente no aumento da necessidade de recursos computacionais para manipular grandes volumes de dados, quando estes são usados em modelos de previsão que simulam a evolução dos fenômenos físicos.

Por outro lado, o estudo dos dados históricos pode ajudar a criar padrões das condições da atmosfera e assim construir um modelo empírico de previsão. Neste trabalho propõe-se o uso de uma rede neural do tipo Perceptron de Múltiplas Camadas (MLP) para construção de um modelo de previsão a partir de um conjunto de dados pré-processado pela TCA.

O pré-processamento é necessário para eliminar redundâncias e buscar a redução da dimensão dos dados para viabilizar o treinamento das redes neurais artificiais, em termos de tempo de processamento para a obtenção de um modelo de previsão climática.

No pré-processamento com a TCA é utilizada uma abordagem de mineração de dados que busca identificar os atributos mais relevantes para o processo de previsão, permitindo assim uma redução da complexidade do problema.

Os dados climatológicos de reanálise utilizados nos experimentos foram extraídos de quatro subáreas do Nordeste do Brasil, onde se procurou aprender o comportamento dos fenômenos meteorológicos.

O restante deste trabalho está organizado da seguinte forma: a Seção 2 faz uma breve introdução ao problema de previsão climática; a Seção 3 traz os princípios teóricos da TCA; a Seção 4 introduz as redes neurais artificiais, destacando o perceptron de múltiplas camadas; a Seção 5 apresenta a metodologia utilizada no trabalho e descreve os dados utilizados; e a Seção 6 traz as conclusões do trabalho.

## 2 Previsão Climática

A previsão climática é definida como a estimativa do comportamento médio da atmosfera com al-

guns meses de antecedência. Por exemplo, em uma escala de tempo sazonal, pode-se prever se o próximo inverno será mais frio que a média, ou ainda, se haverá mais chuva que a estação anterior, ou seja, o objetivo da previsão climática é estimar as propriedades estatísticas do estado climático no futuro (Cavalcanti, 2008).

A previsão do clima realizada por centros de previsão como o CPTEC – Centro de Previsão de Tempo e Clima do Instituto Nacional de Pesquisas Espaciais – INPE ([www.inpe.br](http://www.inpe.br)) e semelhantes, utilizam modelos que procuram descrever o comportamento das condições físico-químicas da atmosfera. Estes modelos são executados em sistemas computacionais de grande porte gerando como saídas possíveis estados futuros da atmosfera com alta resolução. Estas atividades apóiam estudos de impactos e vulnerabilidade e permitem que se façam projeções dos extremos climáticos do estado atmosférico.

### 3 Teoria dos Conjuntos Aproximativos

A Teoria dos Conjuntos Aproximativos (TCA) proposta pelo matemático polonês Zdzislaw Pawlak em 1982, baseia-se nas similaridades entre objetos através da relação de indiscernibilidade. Os dados disponíveis são dispostos em um Sistema de Informação (SI).

#### 3.1 Sistema de Informação

Um Sistema de Informação (SI) é organizado em formato de tabela, em que cada linha representa um objeto e as colunas representam os atributos (Kowowski *et al*, 1999).

Formalmente um SI é um par ordenado  $SI = (U, A)$  em que  $U$  é um conjunto finito de elementos, não vazio chamado de universo, e  $A$  é um conjunto finito, não vazio, de elementos chamados atributos.

Um SI é um Sistema de Decisão (SD) quando inclui um atributo de decisão  $d$ , que não pertence ao conjunto de atributos  $A$ . Formalmente,  $SD = (U, A \cup \{d\})$ , onde  $d \notin A$ .

A Tabela 3.1 apresenta um exemplo de um SD.

Tabela 3.1. Sistema de Decisão.

U	Atributos Condicionais			Decisão
	Mês	Temperatura	Vento	
a <sub>1</sub>	Janeiro	alta	fraco	sim
a <sub>2</sub>	Março	média	moderado	sim
a <sub>3</sub>	Julho	baixa	forte	não
a <sub>4</sub>	Setembro	alta	fraco	não
a <sub>5</sub>	Março	média	moderado	não
a <sub>6</sub>	Janeiro	alta	fraco	sim

O conjunto de objetos é  $U = \{a_1, a_2, a_3, a_4, a_5, a_6\}$ , o conjunto de atributos é  $A = \{\text{Mês, Temperatura, Vento}\}$  e o atributo de decisão  $d = \{\text{Chuva}\}$ .

#### 3.2 Indiscernibilidade

A relação de indiscernibilidade é tomada como uma medida de similaridade entre dois ou mais objetos, ou seja, vários objetos, caracterizados pelos mesmos valores de seus atributos, são indiscerníveis entre si. A relação de indiscernibilidade é definida pela Expressão 3.1 que estabelece que dois objetos  $x$  e  $x'$  do conjunto  $U$ , são indiscerníveis para um subconjunto de atributos  $B \subset A$ , se para cada atributo  $a$ , de  $x$  e  $x'$ , em  $B$ , os valores forem iguais.

$$IND_A(B) = \{(x, x') \in U^2 \mid \forall a \in B, a(x) = a(x')\} \quad (3.1)$$

Considerando o exemplo da Tabela 3.1, para o subconjunto  $B = \{\text{Mês, Temperatura, Vento}\}$ , os objetos  $\{a_1, a_6\}$  e  $\{a_2, a_5\}$  são indiscerníveis, dessa forma, é possível reduzi-los, formando assim duas novas classes consideradas como redutos do conjunto  $U$  (Tabela 3.2). Para a aplicação da TCA neste trabalho, utilizou-se a ferramenta ROSETTA descrita em (Øhrn, 1999).

Tabela 3.2. Redução do Sistema de Decisão.

U	Atributos Condicionais		
	Mês	Temperatura	Vento
C <sub>1</sub>	Janeiro	alta	fraco
C <sub>2</sub>	Março	média	moderado
C <sub>3</sub>	Julho	baixa	forte
C <sub>4</sub>	Setembro	alta	fraco

### 4 Redes Neurais Artificiais

As Redes Neurais Artificiais (RNAs) são técnicas computacionais cujo funcionamento baseia-se na estrutura neural de seres vivos e que realizam tarefas de computação adquirindo conhecimento através de experiência, construída por um processo de aprendizagem. Algumas de suas vantagens importantes são as características de adaptabilidade, generalização e tolerância a falhas (Haykin, 2001).

Existem vários tipos de RNAs que diferem em arquitetura ou forma de treinamento. Para o problema de construção de modelos de previsão, como proposto neste trabalho, que utiliza dados históricos do estado da atmosfera e a informação do fenômeno que se deseja, utiliza-se um Perceptron de Múltiplas Camadas (MLP), com treinamento supervisionado realizado pelo algoritmo de retropropagação do erro (Haykin, 2001).

O algoritmo de retropropagação do erro, baseado nos mínimos quadráticos (regra Delta), consiste na propagação da entrada pelas camadas sucessivas da MLP, até gerar uma saída que é comparada com o

valor real disponível, gerando um erro que é retro-propagado na MLP, camada por camada, sendo seguida pela modificação dos respectivos pesos.

## 5 Metodologia

A área de estudo, para realização dos experimentos, compreende quatro subáreas da região Nordeste do Brasil que são mostradas na Figura 5.1. As coordenadas geográficas são relacionadas na Tabela 5.1, compreendendo 20 pontos de grade para a subárea 1, 15 para a subárea 2, 16 para a subárea 3 e 12 para a subárea 4.

Tabela 5.1. Coordenadas das subáreas de estudo.

Subárea	Longitude	Latitude
1	47°W, 40°W	-7.5°S, 0°
2	40°W, 35°W	-10°S, 0°
3	47°W, 40°W	-17.5°S, -10°S
4	40°W, 35°W	-10°S, -0°

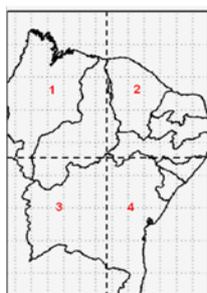


Figura 5.1. Subáreas de estudo.

Em cada subárea os dados são utilizados para construir uma base de dados reduzida, através da aplicação da TCA. Em seguida, a base de dados (ou a redução) obtida forma o conjunto de treinamento da rede neural (MLP) para a obtenção do modelo de previsão da respectiva subárea.

O objetivo do presente trabalho é a obtenção de um modelo de previsão eficiente a partir de um conjunto de dados reduzidos, com base na hipótese de que algumas variáveis são mais adequadas para se estimar o estado futuro da atmosfera. Assim, para efeito de comparação, neste trabalho optou-se pela construção de dois modelos de previsão, baseados em RNA: um gerado a partir do treinamento de uma MLP usando todos os dados disponíveis na base de dados, e outro em que foram usados os dados pré-processados pela TCA, ou seja, os dados com redução da dimensão. O principal objetivo da comparação foi verificar se as reduções podem gerar modelos cujos resultados de previsão são comparáveis com aqueles obtidos utilizando-se todos os dados. A análise de custo computacional não foi tratada neste trabalho. No pré-processamento com a TCA, os atributos relevantes para o processo de previsão climática

são selecionados, como aqueles com maior ocorrência, segundo a relação de indiscernibilidade, formando os redutos desejados para o treinamento da rede neural.

Os dados utilizados nos experimentos conduzidos neste trabalho foram coletados da base de dados de reanálise do NCEP/NCAR (National Centers for Environmental Prediction / The National Center for Atmospheric Research) [<http://www.ncep.noaa.gov/>].

O período de tempo compreende uma janela de 21 anos entre janeiro de 1980 e dezembro de 2000, em uma área contida entre as latitudes [10° N, 35° S] e longitudes [80° W, 30° W], referente à América do Sul. A resolução espacial, em ambas as dimensões da grade, de 2.5° e resolução temporal ( $t$ ) de 1 mês.

As variáveis contidas na base de dados são:

- temperatura do ar (*airt*);
- divergência (*div*);
- precipitação (*prec*);
- umidade específica (*shum*);
- pressão da superfície (*spres*);
- componentes do vento zonal em :
  - 300 hPa (*v300*);
  - 500 hPa (*v500*); e
  - 850 hPa (*v850*)
- componentes do vento meridional em:
  - 300 hPa (*u300*);
  - 500 hPa (*u500*); e
  - 850 hPa (*u850*).

Do conjunto total de dados, foram selecionados 18 anos (janeiro de 1980 a dezembro de 1997) para o treinamento dos modelos de redes neurais e como entrada para o processamento pela TCA. Os demais 3 anos (janeiro de 1998 a dezembro de 2000) foram usados para a validação dos modelos. Destes dados de validação foi selecionado um subconjunto contendo 25% dos dados através de um processo de escolha aleatória, formando o conjunto de teste para o critério de parada antecipada no treinamento das redes.

A métrica para quantificar o desempenho da previsão foi o erro quadrático médio  $E$  dado por:

$$E = \frac{1}{N} \sum_{k=1}^N (y_k - \hat{y}_k)^2 \quad (5.1)$$

em que  $N$  é o número de pontos da grade,  $y_k$  é o valor real no ponto de grade e  $\hat{y}_k$  é a estimativa produzida pela rede. Para a visualização dos resultados obtidos, utilizou-se a ferramenta GrADS (Doty, 2008).

### 5.1 Resultados

Para o processamento das reduções dos atributos utilizou-se o ROSETTA que calcula os redutos mínimos. Os dados são inicialmente discretizados e submetidos a um algoritmo de redução que seleciona os atributos mais relevantes, como aqueles com ocor-

rência igual ou superior a 70% (valor escolhido de forma ad hoc) de presença na função de indiscernimento. Os resultados obtidos neste trabalho são apresentados na Tabela 5.2 para as 4 subáreas de estudo.

Observe-se que das onze variáveis citadas na Seção 5, oito são necessárias nas subáreas 1 e 2, sete na subárea 3 e apenas seis na subárea 4. Estas reduções são utilizadas então para o treinamento das redes neurais de cada subárea, na busca pelo modelo de previsão.

Tabela 5.2. Redutos extraídos pela TCA.

Redutos			
Subárea 1	Subárea 2	Subárea 3	Subárea 4
airt	airt	airt	airt
u850	u850	u500	u500
u500	u500	v850	v850
v850	u300	v500	v300
v500	v500	v300	shum
v300	v300	spres	div
shum	shum	div	
div	spres		

A natureza do problema de previsão de clima, através de uma abordagem inversa como tratada neste trabalho, indica a necessidade de utilizar modelos de redes neurais com aprendizagem supervisionada. Neste trabalho optou-se por utilizar o modelo de rede de perceptrons de múltiplas camadas (MLP), instanciando as redes para as subregiões tratadas, gerando no total 4 redes MLP. Para o treinamento da rede MLP faz-se necessário estipular alguns parâmetros que definem a rede a topologia e o tempo de treinamento para cada rede utilizada.

Para a definição da topologia das redes, cada uma teve sua topologia configurada variando-se, de maneira ad hoc, o número de neurônios nas camadas escondidas e o número máximo de épocas de treinamento, sobre o conjunto de treinamento.

Os testes conduzidos resultaram em uma topologia com uma camada escondida com 10 neurônios, submetida ao máximo de 10000 épocas de treinamento; cada neurônio foi configurado com a função de ativação do tipo logística sigmoideal.

No treinamento de cada rede, foi utilizada a estratégia de parada antecipada com erro calculado sobre o conjunto de teste e comparado com o erro de treinamento durante cada época de treinamento.

Após o treinamento das redes estas foram validadas (generalizadas) utilizando todos os dados e as reduções para gerar previsões sobre os dados de validação. Observa-se que aqui, o conjunto de dados de validação é composto por todos os dados do período 1998 a 2000, incluindo os dados que foram utilizados para teste durante o treinamento.

A Tabela 5.3 apresenta o desempenho dos modelos de previsão, para as quatro estações (est) do ano

de 1998, considerando as 4 subáreas de estudo.

Tabela 5.3. Erro Quadrático Médio.

Erro quadrático médio								
Est.	Subárea 1		Subárea 2		Subárea 3		Subárea 4	
	RNA	TCA	RNA	TCA	RNA	TCA	RNA	TCA
Out.	0,009	0,001	0,421	0,078	0,054	0,585	0,343	0,142
Inv.	0,021	0,002	0,315	0,141	0,348	1,012	1,176	0,176
Pri.	0,043	0,013	0,133	0,180	0,499	2,311	0,547	0,126
Ver.	0,005	0,050	0,441	0,382	0,259	1,763	0,536	0,292

Uma análise qualitativa dos resultados pode ser vista nas Figuras 5.2 a 5.5, que apresenta simultaneamente os resultados obtidos, das quatro estações do ano 1998, nas quatro subáreas analisadas. Os mapas foram gerados utilizando a ferramenta GrADS.

As Figuras 5.2(a), 5.2(b) e 5.2(c) representam respectivamente: as situações observadas (denominadas REAL), a previsão obtida com os dados processados pela TCA e o resultado de previsão com o conjunto de dados completos.

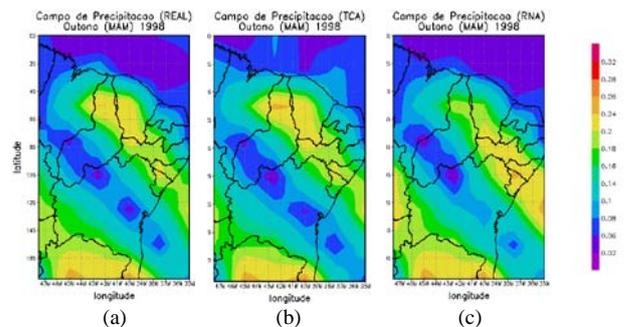


Figura 5.2. Previsão para estação Outono de 1998.

Na Figura 5.3 são apresentados os resultados da estimativa de precipitação para a estação inverno de 1998. Observa-se que a previsão realizada com os dados pré-processados pela TCA tem padrões mais semelhantes àqueles observados na Figura 5.3(a).

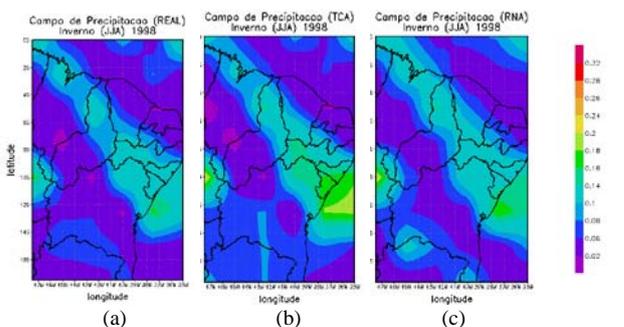


Figura 5.3. Previsão para estação Inverno de 1998.

Na Figura 5.4 são apresentados os resultados de precipitação obtidos para a estação verão de 1998. Observa-se que ambas as estimativas por redes neu-

rais apresentaram padrões muito semelhantes àqueles presentes na Figura 5.4(a).

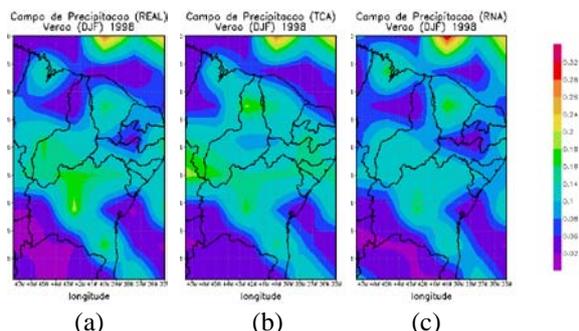


Figura 5.4. Previsão para estação Verão de 1998.

Na Figura 5.5 são mostrados os resultados de precipitação obtidos pelas redes neurais para a estação primavera de 1998. Também observa-se uma grande semelhança entre as estimativas encontradas pelas com a Figura 5.5(a) que representa a informação real desejada.

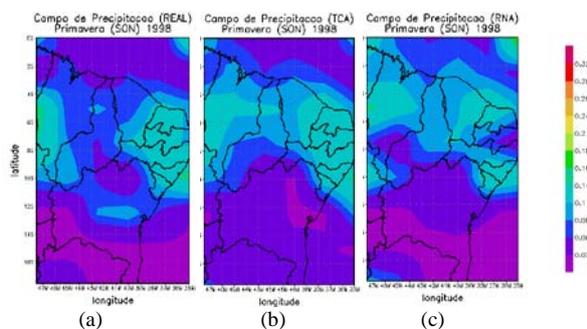


Figura 5.5. Previsão para estação Primavera de 1998.

## 6 Conclusão

Este trabalho apresentou uma proposta de um método de concepção de um modelo de previsão climática utilizando redes neurais.

Dois modelos foram gerados: um utilizando todos os dados disponíveis para treinamento e outro utilizando os dados obtidos de reduções calculadas pela Teoria dos Conjuntos Aproximativos.

Ambos os modelos foram validados com a submissão de um conjunto de dados não utilizado durante o treinamento das redes neurais. Nos testes conduzidos foi possível notar que os conjuntos reduzidos de dados geraram previsões muito compatíveis com as previsões realizadas pelo modelo que utiliza todos os dados, provando que é possível realizar a previsão com base em um subconjunto de variáveis que estão mais relacionadas com o comportamento físico da atmosfera.

Entretanto, como era de se esperar, há uma perda de precisão, como observado na Tabela 5.3 de resultados, sendo este o compromisso associado à simplificação do problema proposta neste trabalho.

Mas, dado que os modelos propostos no trabalho são concebidos utilizando redes neurais artificiais, pode ser possível melhorar a precisão dos resultados com outras estratégias de treinamento das redes neurais.

## Agradecimentos

Os autores agradecem ao INPE e a CAPES pelo apoio para a realização do presente trabalho.

## Referências Bibliográficas

- Cavalcanti, I. F. A. Previsão climática no CPTEC-INPE. (2009). Disponível em: <<http://tucupi.cptec.inpe.br/products/climanalise/cliesp10a/precli.html>>. Acesso em: abr 2008.
- Doty, B. Grid Analysis and Display System (GrADS). Maryland: Center for Ocean-Land-Atmosphere Studies (COLA). Disponível em: <<http://grads.iges.org/grads/head.html>>, Acesso em: 23-fev2009.
- Fayyad U. e Shapiro G.P. e Smyth P. (1996). From Data Mining to Knowledge Discovery in databases. AAAI Press.
- Haykin S. (2001). Redes Neurais: Princípios e Práticas. Bookman, Porto Alegre.
- Komorowski, J. e Øhrn, A. (1999). Modelling prognostic power of cardiac tests using rough sets. Artificial Intelligence in Medicine. pp. 167-191.
- Øhrn A. (1999). Discernibility and Rough Sets in Medicine: Tools and Applications. Tese de Doutorado, Norwegian University of Science and Technology, Department of Computer and Information Science, NTNU.
- Pawlak Z. (1982). Rough sets. International Journal of Computer and Information Sciences, Vol.1. pp. 341-356.
- Vianello, R. L. e Alvez A. R. (2000). Meteorologia Básica e Aplicações, Viçosa, UFV.