Estudo do Comportamento Climático Utilizando uma Abordagem Neuro-Aproximativa

Alex Sandro Aguiar Pessoa
Instituto Nacional de Pesquisas Espaciais
Laboratório Associado de Matemática e Computação Aplicada
asapessoa@lac.inpe.br

José Demisio Simões da Silva
Instituto Nacional de Pesquisas Espaciais
Laboratório Associado de Matemática e Computação Aplicada
demisio@lac.inpe.br

Resumo

A teoria dos conjuntos aproximativos (TCA), trata basicamente da manipulação de informações incertas e redução de dados, tem grande aplicabilidade na mineração devido as suas características de compactação da base de dados e capacidade de extrair informações escondidas. Por isso esta técnica de manipulação de informações foi utilizada de modo conjunto com as redes neurais artificiais.

A tarefa escolhida para testar o uso da metodologia mostrada ao longo deste artigo é por meio de redes neurais estabelecer previsões climáticas para um dado tempo futuro, com níveis de erro aceitáveis, utilizando como pré-processamento dos dados a TCA. A análise por meio da TCA, por sua vez, objetiva a extração das variáveis mais importantes para a região em análise. Após essa análise redes neurais fazem o aprendizado baseado em dados existentes.

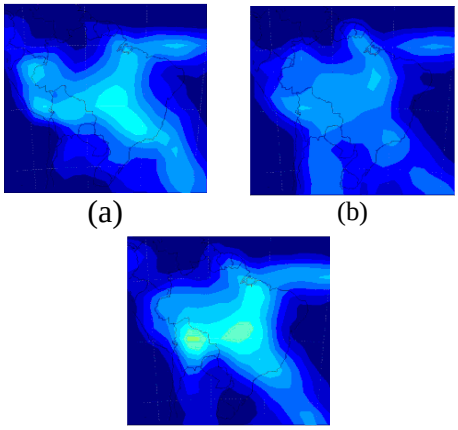
Palavras-chaves: *redes neurais, teoria dos conjuntos aproximativos, previsão climática.*

1. Introdução

A previsão climática no Brasil é realizada através de modelos numéricos que são técnicas que simulam o estado da atmosfera por meio de leis da física, resolvidas numericamente para a realização da previsão.

Neste trabalho, entretanto, busca-se através de métodos utilizados em descoberta de conhecimento (KDD) e na Inteligência Artificial (IA), compreender e aprender sobre o comportamento atmosférico em curto e longo prazo em regiões da América do Sul, com o intuito auxiliar os métodos atuais de previsão climática a fim de diminuir os erros. O aprendizado é realizado através de dados meteorológicos, tais como temperatura, pressão, precipitação.

Embora a atmosfera tenha um comportamento caótico [1] e quando analisada globalmente seja composta de "vários subsistemas" que acabam influenciando uns aos outros, na média, as estações apresentam um comportamento similar, como mostrado nas figuras abaixo:



(a) (b)

(c)

Figura 1 – Precipitação. (a) Jan – 1980; (b) Jan – 1981 e (c) Jan – 1982

Como pode ser visto acima o comportamento da precipitação é parecido, nos anos mostrados. Isto indica que é possível que um sistema computacional aprenda este comportamento, e suas variações no tempo, de acordo com um conjunto de variáveis de entrada.

Para a realização desta tarefa, será mostrado neste trabalho, o emprego da teoria dos conjuntos aproximativos (TCA), para a descoberta das variáveis imprescindíveis, ou de mais importância e a utilização de redes neurais artificiais para o aprendizado propriamente dito.

2. Previsão climática

A previsão climática é uma estimativa do comportamento médio da atmosfera com alguns meses de antecedência. Por exemplo, pode-se prever se o próximo verão será mais quente ou mais frio que o normal, ou ainda, mais ou menos chuvoso. Todavia, tal estimativa não pode dizer exatamente qual será a

quantidade de chuvas ou quantos graus a temperatura estará mais ou menos elevada. Para efeitos didáticos existe uma distinção entre tempo e clima por parte dos meteorologistas. O principal diferencial entre previsão do tempo e previsão clima é a escala temporal. O clima está associado em geral a previsões de longo prazo, como saber a média de temperaturas do próximo verão. Já a previsão do tempo diz respeito à estimativa para pouco tempo, como horas, dias e semanas.

3. KDD

A descoberta de conhecimento em banco de dados (KDD – *Knowledge Discovery in Database*) é o processo de encontrar em um banco de dados padrões que estejam escondidos, sem uma idéia pré-determinada ou hipótese sobre o que são estes padrões [2]. Além de prover esta descoberta de informações a KDD também é composta pelas etapas de: definição dos objetivos, limpeza dos dados, transformação dos dados, mineração de dados e interpretação dos resultados (Figura 2).

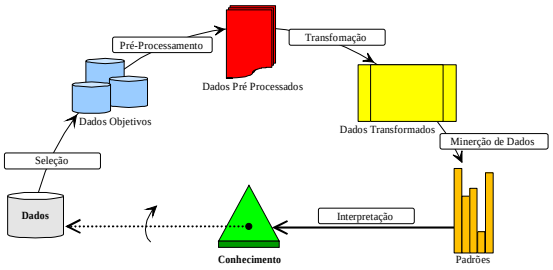


Figura 2: Etapas do processo de KDD

A pesquisa em KDD tem crescido e atraído esforços, baseada na disseminação da tecnologia de bancos de dados e na premissa de que as grandes coleções de dados hoje existentes podem ser fontes de conhecimento útil, que está implicitamente representado e pode ser extraído. No sentido de viabilizar esta tecnologia a KDD se vale, entre outras coisas, de técnicas de aprendizado de máquina, área da inteligência artificial, e de conceitos estatísticos para lidar com a incerteza relacionada às descobertas.

A etapa de maior interesse no processo de descoberta de conhecimento é a mineração de dados, cujo enfoque é a exploração e análise de grandes quantidades de dados para descobrir significativamente modelos e regras. Existem diversas técnicas de mineração de dados, podendo destacar entre tantas: a teoria dos conjuntos aproximativos (*rough sets theory*), teoria dos conjuntos nebulosos (*fuzzy sets theory*), redes neurais artificiais, indução de regras e árvores de decisão.

4. Redes Neurais Artificiais

As redes neurais artificiais (RNA) têm sua inspiração no funcionamento do cérebro, tentando assim imitá-lo por técnicas computacionais com o fim de adquirir, armazenar e utilizar conhecimentos [3]. Assim como no modelo biológico a unidade básica de processamento de uma rede neural é o neurônio. A estrutura topológica a qual estes neurônios estão conectados, em uma RNA é conhecida como arquitetura. Por fim, o processo que adapta a rede a computar uma função desejada é a aprendizagem.

As funções de um neurônio artificial basicamente são:

1. Avaliação dos valores de entrada;
2. Calcular um total para o valor de entradas combinado, ou ponderados pelos pesos w_i;
3. Comparar o total com um limiar (θ);
4. Determinar a saída.

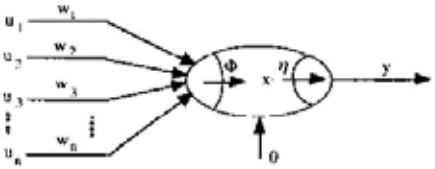


Figura 3: Esquema de um neurônio artificial
Fonte: [http://www.din.uem.br/ia/neurais/]

Nas RNAs, em sua arquitetura tipicamente estão organizadas em camadas, onde estas são classificadas em três camadas:

Camada de entrada: onde os padrões são apresentados à rede.

Camadas intermediárias: onde é feita a maior parte do processamento.

Camada de saída: onde o resultado é apresentado.

As RNAs podem ser classificadas com base na direção na qual o sinal flui [3], podendo ser:

1. *Feedforward:* O sinal se propaga em uma só direção, sendo apresentado na entrada, passando pelas camadas intermediárias e chegando na camada de saída onde são exibidos os resultados.
2. *Feedback:* são conhecidas como redes com realimentação, pois possuem em seu grafo de conectividade ao menos um ciclo.

Existem alguns métodos para a fase de aprendizado de uma RNA, porem o mais utilizado e aplicado neste trabalho é o algoritmo de retro-propagação do erro, conhecido como “backpropagation”. Neste algoritmo a rede opera em dois passos. No primeiro o padrão é apresentado na entrada, flui, então, pelas camadas intermediárias e em seguida o resultado é computado pela camada de saída, onde o valor é comparado ao valor da saída desejada. Se o erro, ocasionado por esta comparação não for o aceitável, então este é propagado da camada de saída até a entrada, para que se corrijam os pesos em função de erro obtido [3].

Este trabalho utiliza em particular uma rede neural do tipo feedforward com o algoritmo de backpropagation.

5. Teoria dos Conjuntos Aproximativos

A TCA é uma extensão da teoria dos conjuntos, cujo enfoque é o tratamento de vagueza e incerteza em dados. Foi inicialmente desenvolvida por *Zdzislaw Pawlak* [10] no início da década de 80, década esta cujo preço e o desempenho dos computadores propiciaram o crescimento e surgimento de novas extensões para a TCA.

5.1. Sistemas de informação

Para o estudo da TCA se faz necessário introduzir alguns conceitos básicos que serão descritos a seguir, começando pela noção de Sistemas de Informação (SI). Um SI nada mais é que um conjunto de dados, sendo representado por uma tabela onde cada linha representa um caso, um evento, um paciente, ou simplesmente um objeto. Toda coluna representa um atributo (uma variável, uma observação, uma propriedade, etc.). Esta tabela, então, é chamada de *sistema de informação*. Mais formalmente, é um par S = (*U*; *A*), onde *U* é um conjunto finito não-vazio de objetos chamado de universo e A é um conjunto finito não-vazio de atributos tal que *a*: $U \rightarrow Va$ para todo $a \in A$. O conjunto *Va* é chamado de conjuntos de valores de *Va*.

Em muitas aplicações, para fins classificatórios, um certo atributo é distinguido e é denominado atributo de decisão. Os sistemas de informação deste tipo são chamados *sistemas de decisão* (SD). Um SD é sistema qualquer de informação na forma S = (*U*; *A*∪{d}), onde d∉*A* é o atributo de decisão. Os elementos de *A* são chamados atributos condicionais ou simplesmente condições [8].

5.2. Indiscernibilidade

Um SD (no formato de uma tabela, por exemplo) pode ser desnecessariamente grande, em pelo menos dois modos: quando elementos "iguais" são representados muitas vezes e quando alguns atributos são supérfluos.

Com relação aos objetos que são representados muitas vezes, existe uma relação de equivalência que tem a capacidade de "tratar" estes problemas de modo que apenas um objeto represente toda uma classe. Esta relação será formalizada abaixo.

Dado S = (*U*; *A*) como sistema de informação, então com qualquer $B \subseteq A$ existe uma relação de equivalência $IND_A(B)$:

$$IND_A(B) = \{(x, x') \in U \mid \forall a \in B, a(x) = a(x')\} \quad (1)$$

$IND_A(B)$ é chamada de relação de *B*-indiscernibilidade. Se $(x, x') \in IND_A(B)$, então objetos *x* e *x'* são indiscerníveis relativamente a qualquer atributo de *B*. A classe de equivalência da relação determinada por *x* pertencente a *X* é denotado $[x]_B$ [8].

5.3. Aproximação dos Conjuntos

A idéia de relação de equivalência induz a partição do universo. Estas partições podem ser usadas para construir novos subconjuntos do universo, que levem em relação alguma característica, ou até mesmo uma determinada classe de um atributo de decisão. Na análise em TCA, levando em conta a relação de indiscernibilidade os elementos podem ser divididos em três classes:

1. Os que podem certamente ser classificados pertencentes a uma desejada classe (não possuem ambigüidades);
2. Os elementos que não podem ser classificados pertencentes a tal classe, pois possuem contradições, como valores iguais para seus atributos e pertencerem a classes diferentes.
3. E os elementos que não pertencem à classe desejada.

Se existem elementos que não podem ser classificados, o conjunto é dito aproximativo (rough set) [7]. Abaixo é mostrada a definição formal das noções apresentadas:

Dado S = (U; A), sendo um sistema de informação e $B \subseteq A$ e $X \subseteq U$. Nós podemos aproximar *X* usando somente as informações contidas em *B* construindo as aproximações *B*-inferiores e *B*-superiores de *X*, denotados respectivamente $\underline{B}X$ e $\overline{B}X$, onde:

$$\underline{B}X = \{x \mid [x]_B \subseteq X\} \text{ e } \overline{B}X = \{x \mid [x]_B \cap X \neq \varnothing\} \quad (2) \text{ e } (3)$$

Os objetos em $\underline{B}X$ podem ser certamente classificados como membros de *X* na base de conhecimento *B*, enquanto os objetos em $\overline{B}X$ podem somente serem classificados como possíveis membros de *X* na base de conhecimento *B*. O conjunto $F_B(X) = \underline{B}X - \overline{B}X$ é chamada de região de *B*-fronteira de *X*, sendo que estes objetos não podem ser classificados pertencentes a *X* na base de conhecimento *B* com absoluta certeza. O conjunto $E_B(X) = U - \overline{B}X$ é então chamado de *B*-região externa de *X*, e estes objetos certamente podem ser classificados como não pertencentes a *X* (na base de conhecimento *B*).

Portanto um conjunto é dito *aproximativo* se a região da fronteira não é vazia, e será chamado de *preciso* (crisp) caso contrário.

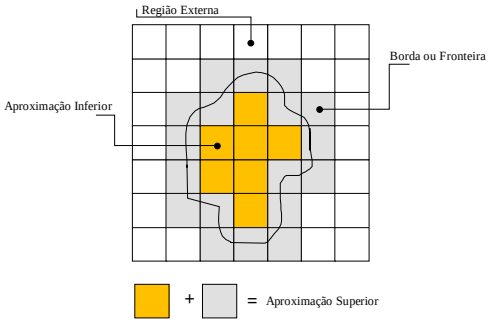


Figura 4. Aproximações dos Conjuntos

5.4. Reduções

Anteriormente foi citado que os modos de reduzir a base de dados para análise era de representar elementos iguais com um só registro, através da relação de indiscernibilidade, e o outro artifício para redução é manter somente os atributos que preservam a relação de indiscernibilidade, ou seja, eliminar os supérfluos. Normalmente existem vários subconjuntos de atributos que possuem esta característica e os que são mínimos são chamados de *reduções*. A determinação das reduções é um problema "*NP-hard*" [14]. O número de reduções de um SI com m atributos podem ser iguais a:

$$\binom{m}{\lfloor m/2 \rfloor} \quad (4)$$

O cálculo das reduções é uma tarefa cujo custo computacional é muito alto, fazendo com que este seja o maior gargalo da TCA, exigindo esforços para desenvolvimentos de novos algoritmos e utilização de heurísticas que tornem a computação mais rápida.

Dado o sistema de informação S = (*U; A*) as definições das noções apresentadas até então são mostradas a seguir:

Uma redução de S é um conjunto mínimo de atributos $B \subseteq A$ tal que $IND_S(B) = IND_S(A)$. Em outras palavras, uma redução, RED(*B*), é o conjunto mínimo de atributos de *A* que preserva o particionamento do universo realizado pela ralação de indiscernibilidade, e conseqüentemente a habilidade para executar classificações assim como o conjunto de atributos (completo) a faz. O núcleo pode ser dado por $N(B) = \cap RED_i(B)$, i = 1...n.

Sendo S um SI com *n* objetos, a matriz de discernibilidade de S é uma matriz simétrica *n* x *n* com entradas c_{ij}, que é dado na equação abaixo. Cada entrada consiste em um conjunto de atributos que difere os objetos x_i e x_j.

$$c_{ij} = \{a \in A \mid a(x_i) \neq a(x_j)\} \text{ para } i,j = 1, ..., n \quad (5)$$

A função de discernibilidade f_A para um sistema de informação S é uma função de m variáveis Booleanas: $a_1^*,...,a_m^*$ (correspondente aos atributos a1,..., am), cuja definição é dada abaixo e onde $c_{ij}^* = \{a^* \mid a \in c_{ij}\}$:

$$f_A(a_1^*,...a_n^*) = \wedge\{\vee\ c_{ij}^* \mid 1 \leq j \leq i \leq n,\ \ c_{ij} \neq \varnothing\} \quad (6)$$

O conjunto de todos implicantes primos de f_A determina o conjunto de todas as reduções de *A*.

Cada linha da função de discernibilidade corresponde a uma coluna da matriz de discernibilidade. Esta matriz é simétrica e com a diagonal vazia.

5.5. Discretização de Variáveis Meteorológicas

O processo de discretização consiste basicamente em transformar variáveis não categóricas em variáveis categóricas. Este processo faz com que o mundo seja visto de uma forma mais grosseira, porém determinadas metodologias, como no caso da TCA, trabalham melhor com variáveis numéricas discretizadas do que contínuas isto porque uma das premissas da TCA é reduzir a precisão dos dados revelando suas regularidades [11].

Existem diversos métodos para realização desta tarefa, contudo no caso de variáveis meteorológicas há um interesse em apenas três classes (Figura 5):

1. Abaixo da média histórica;
2. Na média histórica;
3. Acima da média histórica.

Então de acordo com a necessidade usou-se seguinte técnica para encontrar estas classes:

Sendo $X=\{x_i \in \mathbf{R}\}$, i=1..n; n é o número de elementos; θ = limiar do intervalo (θ ∈ [0, 1])

a = mínimo(X);
b = máximo(X);

$$P_m = \frac{\sum_{i=1}^{n} x_i}{n}; \quad (7)$$

$$a' = a + ((P_m - a) * \theta); \quad (8)$$

$$b' = P_m + ((b - P_m) * \theta); \quad (9)$$

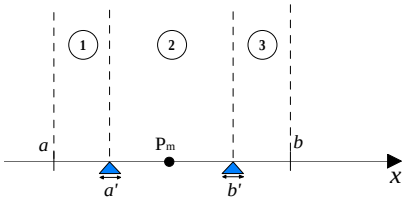


Figura 5 – Discretização

Deste modo as variáveis são discretizadas nas três categorias e estas ainda podendo ser "alargadas" ou "estreitadas" de acordo como o parâmetro θ.

5.6. Rosetta

O sistema ROSETTA (*Rough Set Toolkit for Analysis of Data*) é uma ferramenta baseada na TCA, que cobre as diversas etapas do processo de KDD [9], o que possibilita diferentes abordagens e modelagens. Mas uma vantagem de extrema importância do ROSETTA é a possibilidade de inclusão de novos algoritmos, caso haja necessidade, pois este é um sistema modular, ou seja, construído em blocos.

6. Metodologia

As análises foram feitas para a previsão de precipitação de 1 e 3 estações adiante (Δ). Para a realização das análises foram coletados dados do NOAA em [http://www.cdc.noaa.gov/], para o período de janeiro de 1980 a dezembro de 2000 e a área contida entre as latitudes [10° N, 35° S] e longitudes [80° W, 30° W], referente à América do Sul, em uma resolução espacial em ambas as dimensões (x, y) de 2.5° e resolução temporal (t) de 1 mês. As variáveis coletadas são mostradas abaixo:

	Variável	Descrição	Unidade
1	airt	Temperatura do ar	°C
2	div	divergência	1/s
3	estacao	Estação do ano	–
4	lat	Latitude	graus
5	lon	Longitude	graus
6	precAtual	Precipitação Atual	mm/dia
7	shum	Umidade Específica	kg/kg
8	spres	Pressão da Superfície	mb
9	temp	Temperatura	°C
10	u300	Vento Zonal 300 hPa	m/s
11	u500	Vento Zonal 500 hPa	m/s
12	u850	Vento Zonal 850 hPa	m/s
13	v300	Vento Meridional 300 hPa	m/s
14	v500	Vento Meridional 500 hPa	m/s
15	v850	Vento Meridional 850 hPa	m/s

Adicionalmente as variáveis mostradas acima foram incorporadas nas análises as componentes espaciais e temporais (longitude, latitude e tempo).

Para simplificação e agilidade computacional serão analisadas cinco subáreas da América do Sul, dispostas a cobrirem uma parte das cinco regiões brasileiras, pois devido a grande extensão territorial do Brasil, há muitos regimes de precipitação e conseqüentemente cada uma com seu clima típico.

As regiões são (Figura 6):

1. Norte (N):
long = 67.5°W, 57.5°W; lat = -7.5°S, 0°;

2. Nordeste (NE):
long = 45°W, 35°W; lat = -7.5°S, 0°;

3. Centro-Oeste (CO):
long = 62.5°W, 52.5°W; lat = -22.5°S, -15°S;

4. Sudeste (SE):
long = 52.5°W, 42.5°W; lat = -27°S,.5 -20°S;

5. Sul (S):
long = 60°W, 50°W; lat = -35°S, -27.5°S;

Figura 6 – Regiões Selecionadas para análise

A análise é constituída de três fases:

1. Calcular as reduções do SI para a região correspondente, no sistema Rosetta. Nesta fase são realizadas todas as etapas do processo de KDD (Figura 2). O resultado é um conjunto de reduções onde as variáveis com maior ocorrência são escolhidas como reduções para o treinamento da rede neural.
2. Treinar a rede neural com o SI sem reduções de variáveis, ou seja, com todas as variáveis mostradas anteriormente;
3. Treinar a rede neural com o SI com reduções de variáveis. A partir das reduções provenientes da fase de KDD, a rede é treinada e por fim os resultados são comparados com a rede neural de conjunto de variáveis completa.

7. Resultados

A. Escolha das variáveis (Processo de KDD)

Nesta fase as dados foram pré-processados com o intuito de discretizar as variáveis de acordo com a seção 5.5. A seguir, no sistema Rosetta foram computadas as reduções para cada região em análise.

Com base no conjunto total de reduções foram selecionadas as variáveis com uma ocorrência igual ou superior a 70%. As variáveis para Δ = 1 e Δ = 3 são:

Tabela 1: Variáveis com ocorrência ≥70%

Δ = 1				
CO (289)	**N (274)**	**NE (277)**	**S (313)**	**SE (279)**
Variável	**Variável**	**Variável**	**Variável**	**Variável**
u300	**shum**	**precAtual**	**airt**	**airt**
u850	**precAtual**	**shum**	**lat**	**u300**
precAtual	**estacao**	**temp**	**estacao**	**temp**
u500	**lat**	**lon**	**lon**	**estacao**
lat	**lon**	**lat**		**lat**
lon		**estacao**		**lon**
estacao				

Tabela 2: Variáveis com ocorrência ≥ 70%

Δ = 3				
CO (306)	**N (292)**	**NE (276)**	**S (313)**	**SE (270)**
Variável	**Variável**	**Variável**	**Variável**	**Variável**
u850	precAtual	precAtual	airt	temp
u500	lon	temp	estacao	lat
u300	lat	lat	lat	estacao
spres	estacao	estacao	lon	
lat		lon		
lon				
estacao				

Obs: Os números entre parênteses são o número total de reduções.

B. Treinamento e avaliação das RNA

As redes neurais nesta etapa são apresentadas a um conjunto de dados de treinamento de 18 anos, contados a partir de janeiro de 1980. A principio os dados são apresentados com todas as variáveis a rede neural, com o objetivo de atingir um erro de 5% do valor máximo da massa de dados ou 1000 iterações. A seguir um conjunto de dados do mesmo período que o anterior, porém com o número de variáveis reduzido, de acordo com a região em questão é processada em outra RNA. Os resultados são exibidos abaixo (EQM – erro quadrático médio):

Tabela 3: EQM - Treinamento de RNA (Δ = 1)

	Δ = 1		Δ = 3	
	Completo	**Reduzido**	**Completo**	**Reduzido**
CO	0.6119	0.6557	0.8240	0.7290
N	0.7662	0.9924	0.8400	1.0722
NE	1.0471	1.8625	1.6960	2.2432
S	0.8321	1.3546	0.8180	1.2736
SE	0.8361	0.8928	0.6500	1.0179

Para a avaliação foram utilizados os três últimos anos dos dados coletados, com a data inicial em janeiro de 1998. Os dados são apresentados às respectivas redes, ou seja, dados sem redução na rede neural treinada sem redução e assim por diante.

Tabela 4: EQM - Avaliação de RNA (Δ = 1)

	Δ = 1		Δ = 3	
	Completo	**Reduzido**	**Completo**	**Reduzido**
CO	0.96732	0.55592	0.82986	0.44119
N	2.1234	1.8113	2.395	2.4074
NE	3.0413	1.9976	3.3929	1.4138
S	1.3943	1.2811	1.4813	1.0142
SE	0.93962	0.8339	1.3789	0.86613

8. Conclusões

Como pode ser observado nas Tabela 6 e 7 os dados treinados em uma RNA com reduções no número de variáveis em quase todos os casos apresentou o erro quadrático médio (EQM) menor do que a versão da rede treinada com todas as variáveis.

Vale ressaltar, que a previsão climática com redes neurais, para fim metodológico, foi realizada para prever a precipitação e em algumas regiões da América do Sul, mas seria possível gerar previsões para outras variáveis e regiões maiores.

Referências

[1] Cavalcanti, I. F. A. et al. *Global Climatological Features in a Simulation Using the CPTEC-COLA AGCM*. Jounal of Climate, vol. 15, n 27, p. 2965-2988. 2002

[2] Chen, Z. *Data Mining and Uncertain Reasoning: an Integrated Approach*., John Wiley & Sons, New York 2001.

[3] Fernandes, A. M. da R. Inteligência Artificial: Noções Gerais, Visual Books, Florianópolis, 2003.

[4] Holsheimer M. & Siebes A. *Data mining: the search for knowledge in databases*. Report CSR9406. Amsterdam, the Netherlands: CWI, Jan. 1994.

[5] Lorenz, E. N. *Deterministic non-periodic flow*. J. Atmos. Sci., v. 20, p. 130-141, 1963.

[6] Lorenz, E. N. *A study of the predictability of a 28-variable atmospheric model*. Tellus, v. 17, p. 321-333, 1965.

[7] Lorenz, E. N. *The predictability of a flow which possesses many scales of motion*. Tellus, v. 21, p. 289-307,1969.

[8] Komorowski, J. *et al*. *Rough sets: A tutorial*. in: S.K. Pal and A. Skowron (eds.), Rough fuzzy hybridization: A new trend in decision-making, Springer-Verlag, Singapore, 1999.

[9] Øhrn, A. *Discernibility and Rough Sets in Medicine: Tools and Applications*. PhD thesis, Norwegian University of Science and Technology, Department of Computer and Information Science, NTNU. 1999.

[10] Pawlak Z. *Rough sets*. International Journal of Computer and Information Sciences, 11:341--356, 1982.

[11] Pawlak, Z. *Rough sets – Theorical aspects of reasoning about data*, Kluver Academic Publishers, Dordrecht, 1991.

[12] Pawlak, Z.; SKOWRON, A. *Rough membership functions*. In: R. Yager, M. Fedrizzi J. Kacprzyk (Eds.), Advances in the Dempster-Shafer Theory of Evidence, Wiley, NewYork, p. 251-271, 1994.

[13] Røed G. *Knowledge Extraction from Process Data: A Rough Set Approach to Data Mining on Time Series*, "citeseer.nj.nec.com/119626.html", 1999

[14] Skowron, A. & Rauszer C. *The Discernibility Matrices and Functions in Information Systems*. In: Slowinski, p. 331 – 362, 1992.

[15] Wong, S. K. M. & Ziarko, W. *Comparison of the probabilistic approximate classification and the fuzzy set model*. Fuzzy Sets and Systems 21, p. 357-362, 1986.