

UTILIZANDO UMA REDE NEURAL ARTIFICIAL PARA APROXIMAÇÃO DA FUNÇÃO DE EVOLUÇÃO DO SISTEMA DE LORENTZ

A.Martiniano^{1*}; R.P.Ferreira¹; A.Ferreira²; A.Ferreira³; R.J.Sassi¹

1 Universidade Nove de Julho, 01504-001, São Paulo-SP, Brasil

2 Universidade de São Paulo - USP, 05508-900, São Paulo-SP, Brasil

3 Faculdade Santa Rita de Cássia, 02273-001, São Paulo-SP, Brasil

*andrea.martiniano@gmail.com

Artigo submetido em 07/09/2015 e aceito em 25/02/2016

RESUMO

O objetivo principal deste artigo é realizar a aproximação da função de evolução temporal do Sistema de Lorenz utilizando uma Rede Neural Artificial do tipo MLP (Multilayer Perceptron). Além deste objetivo principal, como objetivo específico, apresentam-se os conceitos básicos das Redes Neurais Artificiais (RNAs), um breve histórico da Teoria do Caos e o Sistema de Lorenz. A metodologia adotada na estruturação deste artigo foi definida como bibliográfica e experimental. Atualmente, existe grande interesse nos modelos de redes neurais para resolver problemas não convencionais e complexos, nesse

contexto, as Redes Neurais Artificiais têm surgido como alternativa para inúmeras aplicações em diversas áreas do conhecimento. Os resultados obtidos nos experimentos apontam positivamente para a utilização das RNAs. Espera-se com esse artigo incentivar a utilização das RNAs em aplicações complexas em que a aprendizagem, associação, generalização e abstração são necessárias para apoio à tomada de decisão. Concluiu-se que a utilização das RNAs pode ser uma alternativa para resolução de problemas envolvendo aproximação de funções.

PALAVRAS-CHAVE: Rede Neural Artificial, Sistema de Lorenz, Aproximação de Função.

USING AN ARTIFICIAL NEURAL NETWORK TO APPROXIMATE THE TEMPORAL EVOLUTION FUNCTION OF THE LORENZ SYSTEM

ABSTRACT

The main objective of this paper is to approximate the temporal evolution function of the Lorenz system using Artificial Neural Networks (ANN) type Multilayer Perceptron (MLP). Apart from this main objective, as a specific objective, presents the basic concepts of ANN, a brief history of chaos theory and the Lorenz system. The methodology used in the structuring of this paper was defined as bibliographic and experimental. Currently, there is great interest in models of neural networks to solve unconventional and complex problems, in this

context the ANN have emerged as an alternative for numerous applications in various areas of knowledge. The results of the experiments indicate positively to the use of ANN. It is hoped that this paper encourage the use of ANN in complex applications where learning, association, generalization and abstraction are needed to support decision-making. It was concluded that the use of ANN could be an alternative for solving problems involving approximation functions.

KEYWORDS: Artificial Neural Network, Lorenz system, Approximation functions.

1. INTRODUÇÃO

A Teoria do Caos estabelece que uma pequena mudança ocorrida no início de um evento qualquer pode ter consequências desconhecidas no futuro. Isto é, uma ação nesse exato momento, terá um resultado posterior, embora desconhecido. A Teoria do Caos é um padrão de organização dentro de um fenômeno desorganizado, ou seja, dentro de uma aparente casualidade (TEORIA DO CAOS, 2016).

O meteorologista norte-americano Edward Norton Lorenz (1917-2008) descobriu, no início da década de 1960, que acontecimentos simples tinham um comportamento tão desordenado quanto à vida. Ele chegou a essa conclusão após testar um programa de computador que simulava o movimento de massas de ar.

Em busca de uma resposta Lorenz inseriu um dos números que alimentavam os cálculos da máquina com algumas casas decimais a menos, na expectativa de que o resultado tivesse poucas variações. No entanto, a pequena alteração transformou completamente o padrão das massas de ar. Segundo ele seria como se o bater das asas de uma borboleta no Brasil causasse, tempos depois, um tornado no Texas. Fundamentado em seus estudos, ele formulou equações que demonstravam o Efeito Borboleta (MONTEIRO, 2006).

Origina-se assim a Teoria do Caos. Alguns cientistas concluíram também que a mesma imprevisibilidade aparecia em quase tudo, do número de vezes que o olho pisca até o Comportamento do Tráfego Urbano ou a Cotação da Bolsa de Valores. Para reforçar essa teoria, na década de 1970 o matemático polonês Benoit Mandelbrot notou que as equações de Lorenz coincidiram com as que ele próprio havia feito quando desenvolveu os fractais.

As Redes Neurais Artificiais (RNAs) são modelos constituídos por unidades de processamento simples, chamados de neurônios artificiais, que calculam funções matemáticas. Estes modelos são inspirados na estrutura do cérebro e têm como objetivo simular o comportamento humano, tais como: aprendizagem, associação, generalização e abstração quando submetidas a treinamento. As RNAs são particularmente eficientes para o mapeamento entrada/saída de sistemas não lineares e para realizar processamento paralelo, além de simular sistemas complexos (HAYKIN, 2001).

O artigo está organizado após essa seção introdutória da seguinte forma: a seção 2 são apresentados os materiais e métodos, na seção 3 é apresentada os conceitos básicos das Redes Neurais Artificiais, na seção 4 é apresentado o Sistema de Lorenz, na seção 5 são apresentados os Resultados e Discussão e na seção 6 o artigo é encerrado com as Considerações Finais.

2. MATERIAIS E MÉTODOS

A realização da pesquisa deste artigo está embasada em consultas às fontes bibliográficas e de referencial teórico como: artigos, livros, teses, dissertações e websites com conteúdo sobre Teoria do Caos, Efeito Borboleta e Redes Neurais Artificiais.

Utilizou-se, neste artigo, uma Rede Neural Artificial (RNA) do tipo MLP (Multilayer Perceptron). A plataforma de hardware utilizada nos experimentos foi um computador com processador Intel® Core™ 2 Duo CPU T6600 de 2,20 GHz com 4,00 GB de memória RAM DDR3, 320 GB de disco rígido e sistema operacional de 64 Bits.

Com relação ao software utilizado na Rede Neural Artificial do tipo MLP, foram usadas as sub-rotinas disponíveis no software MATLAB® version 7.8.0 R2009a. As RNAs são consideradas um aproximador universal de funções contínuas (KOVÁCS, 2006).

Segundo Russel e Norvig (2013), aprender a aproximar funções é considerado uma tarefa de inferência indutiva, a RNA aprende a representar funções de forma eficiente e útil. As características das RNAs confirmam a sua aplicação em diversas áreas como: previsão de falência (ITURRIAGA e SANZ, 2015), previsão da volatilidade dos preços do ouro (KRISTJANPOLLER e MINUTOLO, 2015), previsão de máxima e mínima do preço das ações de empresas de distribuição de energia elétrica (LABOISSIERE, FERNANDES e LAGE, 2015), análise de desempenho do mercado acionário (RATH, JAGADEV e NAYAK, 2015), estimativas de vendas (LARRAIN e TURNER, 2014), previsão do absenteísmo no trabalho (MARTINIANO et al., 2012), previsão do comportamento do tráfego veicular urbano (FERREIRA, 2011), identificação de demonstrações financeiras falsificadas (GAGANIS, 2009).

Outras aplicações também podem ser encontradas em campos diversos como: modelagem, análise de séries temporais, reconhecimento de padrões e processamento de sinais e controle (HAYKIN, 2001; SILVA, SPATTI e FLAUZINO, 2010; BRAGA, CARVALHO e LUDERMIR, 2011).

3. REDES NEURAIS ARTIFICIAIS

As pesquisas sobre as RNAs tiveram início em 1943, com o artigo de Warren McCulloch, fisiologista e psiquiatra, e Walter Pitts, matemático. Em geral, aceita-se como o início do estudo sobre RNAs o artigo de McCulloch e Pitts (1943). Neste artigo é proposto que, com um número suficiente de neurônios artificiais, uma rede seja dotada de capacidade de realizar uma aproximação de qualquer função contínua.

Pode-se destacar adicionalmente os trabalhos de Hebb (1949) e Rosenblatt (1958) por suas contribuições no desenvolvimento da teoria sobre RNAs.

As RNAs são modelos inspirados na estrutura do cérebro tendo como objetivo simular o comportamento humano em processos como: aprendizagem, adaptação, associação, tolerância a falhas, generalização e abstração (HAYKIN, 2001; SIMÕES e SHAW, 2007; SILVA, SPATTI e FLAUZINO, 2010).

Nestas redes, a aprendizagem se dá através de um conjunto de unidades simples de processamento chamados de neurônios artificiais ou nós. Na Figura 1 é mostrada uma representação do neurônio artificial.

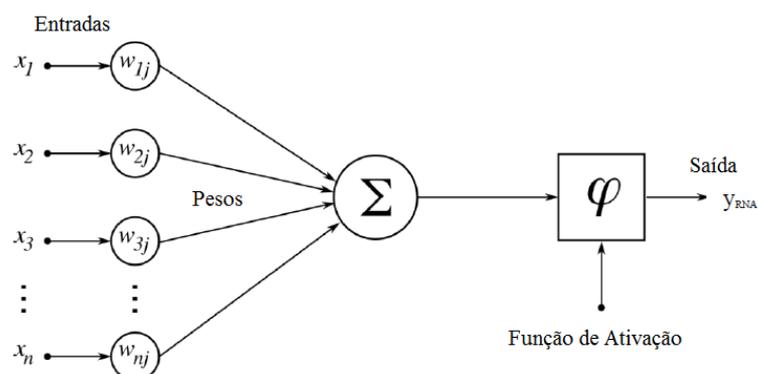


Figura 1 – Representação do neurônio artificial.

Fonte: Adaptado de Affonso (2010).

Observa-se na Figura 1 os dados (vetores de dados) de entrada da rede (X_1, \dots, X_n), os neurônios da camada de entrada da rede (W_{1j}, \dots, W_{nj}) com seus respectivos pesos, logo em seguida a junção aditiva ou soma representada pela letra sigma, em seguida a função de ativação e finalmente a camada de saída (Y_{RNA}), formada por um neurônio.

Uma importante característica das RNAs é a sua capacidade de aprender a partir de dados incompletos e sujeitos a ruídos. Em um sistema computacional convencional, se uma parte falha, em geral, o sistema como um todo se deteriora, enquanto que, em uma RNA, a tolerância às falhas faz parte da arquitetura, devido à sua natureza distribuída de processamento. Se um neurônio falha, sua saída errônea é sobrescrita pelas saídas corretas de seus elementos vizinhos. Assim, a princípio, uma RNA exibe uma degradação suave do desempenho em vez de apresentar uma falha catastrófica (HAYKIN, 2001). Uma das primeiras RNAs, cuja arquitetura foi baseada no neurônio biológico, foi proposta por Rosenblatt (1958): o *Perceptron*. O objetivo desta rede é classificar as entradas x_i (ou estímulos) em duas classes através de um hiperplano. Para o caso simples de um espaço em duas dimensões, o hiperplano fica reduzido a uma reta, cuja Equação é representada em (1):

$$\sum x_i w_i + w_0 = 0 \quad (1)$$

A ativação do neurônio artificial é realizada através da função de ativação, ϕ a qual desempenha tarefa semelhante à sinapse no neurônio biológico, transmitindo ou bloqueando os impulsos nervosos. De uma forma geral, a aprendizagem das redes *Perceptron* se dá através dos ajustes dos pesos sinápticos. O valor do peso sináptico $w_i^{(t+1)}$ no instante $t+1$, será determinado em função do seu valor na iteração anterior w_i^t , conforme a Equação a seguir:

$$w_i^{t+1} = w_i^t + \Delta w_i^t \quad (2)$$

A atualização dos pesos depende do algoritmo, mas geralmente baseia-se na minimização do erro ε_i , entre os valores previstos pela rede e as saídas y_i desejadas, conforme Equação (3):

$$\varepsilon_i = \sum w_i x_i - y_i \quad (3)$$

Desta forma, o aprendizado (ou treinamento) em uma RNA é definido como o ajuste iterativo dos pesos sinápticos, de forma a minimizar os erros (HAYKIN, 2001).

Uma definição geral do que vem a ser aprendizado em uma RNA pode ser expressa da seguinte maneira: “Aprendizado é o processo pelo qual os parâmetros de uma RNA são ajustados através de uma forma continuada de estímulo pelo ambiente no qual a rede está operando, sendo o tipo específico de aprendizagem realizada definido pela maneira particular como ocorrem os ajustes realizados nos parâmetros” (MENDEL e MCLAREN, 1970).

Diversos métodos para aprendizado foram desenvolvidos, podendo ser agrupados em dois paradigmas principais: aprendizado supervisionado e aprendizado não-supervisionado.

No aprendizado supervisionado, existe um conhecimento prévio sobre os valores das entradas x_i e respectivas saídas y_i . A este conjunto de pares ordenados (x_i, y_i) , que se conhece a priori, dá-se o nome de base de dados de aprendizado. O algoritmo mais difundido é o de retro propagação do erro (*error back-propagation*) utilizado pela RNA do tipo *MultiLayer Perceptron* (MLP) utilizada neste artigo.

O algoritmo de treinamento *error back-propagation* funciona da seguinte maneira: apresenta-se um padrão à camada de entrada da rede, este padrão é processado, camada por camada, até que a saída forneça a resposta processada, f_{MLP} , calculada como mostrado a seguir, na Equação (4). Onde v_l e w_{lj} são pesos sinápticos; b_{l0} e b_0 são os *biases*; e φ a função de ativação.

$$f_{MLP}(x) = \varphi \left(\sum_1^{Non} v_l \cdot \varphi \left(\sum w_{lj} x_l + b_{l0} \right) + b_0 \right) \quad (4)$$

Na aprendizagem não supervisionada não se conhece os valores das saídas desejadas y_i , então a aprendizagem ocorre através da identificação de padrões nas entradas. O aprendizado de uma RNA, na maioria dos casos, acontece com um subconjunto de exemplos (vetores de dados) que definem o chamado conjunto de treinamento e o teste da RNA é realizado com outro subconjunto de exemplos (vetores de dados) que definem o chamado conjunto de teste.

A arquitetura de uma RNA depende do tipo de problema no qual a rede deve ser utilizada, pois ela é definida, entre outros fatores, pelo número de camadas, número de nós em cada camada, pelo tipo de conexão entre os nós (*feedforward* ou *feedback*) e por sua topologia (Recorrentes e Não-Recorrentes) (KOVÁCS, 2006). A Tabela 1 apresenta alguns parâmetros adotados no aprendizado das RNAs.

Tabela 1 – Parâmetros de aprendizado (treinamento) para arquiteturas de rede neural artificial.

Parâmetro	Arquiteturas de RNAs	Função do Parâmetro
Taxa de Aprendizado	Todas	Controla a velocidade do ajuste dos pesos da rede neural artificial.
Momentum	<i>Error Back-Propagation</i> com MLP	Acompanha passo a passo o parâmetro taxa de aprendizado. Sua função é monitorar a possível ocorrência de oscilações grandes nos valores nas conexões de peso.
Tolerância de Erro	<i>Error Back-Propagation</i> com MLP	É utilizado para especificar a máxima diferença entre a saída desejada e saída gerada pela RNA. Este parâmetro especifica quão próximo a saída da RNA deve estar da saída desejada.
Função de Ativação	Todas	Seleciona a função de ativação, a qual é utilizada pela unidade de processamento da rede neural artificial. Podem ser utilizadas as funções <i>sigmoide</i> , tangente hiperbólica, função <i>gaussiana</i> , etc.
Vizinhança	Rede SOM	Na rede SOM, uma área é definida ao redor do neurônio vencedor, onde os neurônios não vencedores têm seus pesos alterados em algum grau.
Número de Épocas	Rede SOM, MLP	O número de épocas determina o número de passos para o treinamento da rede através do treinamento dos dados. Uma época ocorre ao final da apresentação de todos os itens de dados exatamente uma vez.

Fonte: Adaptado de Bigus (1996) e Sassi (2006).

As RNAs podem ser treinadas utilizando-se valores iniciais aleatórios para as conexões de pesos. Os parâmetros de aprendizado são inicializados e os padrões de treinamento dos vetores de dados são apresentados para a RNA. Ao longo do progresso do treinamento são ajustadas as conexões de pesos e é possível monitorar a *performance* da RNA.

Atualmente, existe grande interesse nos modelos de redes neurais para resolver problemas não convencionais, nos últimos anos às redes neurais artificiais têm surgido como alternativa para inúmeras aplicações. Nesta seção foram abordados os conceitos básicos das RNA's.

4. O SISTEMA DE LORENZ (SISTEMA DINÂMICO)

Nesta seção apresenta-se um breve histórico do Sistema de Lorentz.

Edward Norton Lorenz (1917-2008) tornou-se professor de ciências atmosféricas do MIT em 1962. Sua formação original foi em Matemática. Mas, após servir durante a Segunda Guerra como meteorologista, resolveu continuar nessa área e trabalhar com aspectos teóricos. Como todo meteorologista, ele estava interessado no problema da previsão do tempo. E como todo ser humano, ele se perguntava: por que não se consegue prever o tempo que fará nos próximos dias, com um grau de confiabilidade razoável, apesar de se conhecer as equações da circulação atmosférica e as condições climáticas: temperatura, pressão, umidade, velocidade do vento etc. (MONTEIRO, 2006).

Na época, usava-se um modelo linear para previsão meteorológica, baseado em dados experimentais. Com esse modelo, a temperatura do dia seguinte era prevista como sendo uma constante C_0 , mais C_1 vezes a temperatura de hoje ao meio-dia, mais C_2 vezes a umidade relativa do ar na cidade vizinha às 11h00min, mais C_3 vezes a velocidade do vento em outra cidade às 15h00min, etc, ou seja, através de uma combinação linear dos dados obtidos. Havia procedimentos matemáticos para estimar os valores convenientes das constantes C_j , a partir dos dados colhidos no presente e no passado. O papel do meteorologista era selecionar os termos a serem multiplicados por essas constantes. N. Wiener (1894-1964) acreditava que esse método linear bastava para solucionar satisfatoriamente qualquer problema relacionado com a previsão de uma série temporal. Lorenz, entretanto, não estava convencido disso.

Em 1956, Lorenz decidiu testar a crença de Wiener do seguinte modo. Ele selecionaria um sistema de equações diferenciais não lineares e o resolveria num computador, obtendo, assim, uma solução numérica. Essa solução seria tratada como se fosse uma coleção de dados reais e seria usada para avaliar a precisão do método linear. Lorenz resolveu simplificar bastante as equações da circulação atmosférica, a fim de obter uma solução numérica rápida com o computador que dispunha capaz de realizar, por segundo, 400 adições ou 60 multiplicações (MONTEIRO, 2006).

As equações que descrevem a dinâmica dos processos atmosféricos são: a Equação de Navier-Stokes:

$$\rho \left(\frac{\partial \vec{v}}{\partial t} + \vec{v} \cdot \nabla \vec{v} \right) = \vec{F} + \vec{F}_c - \nabla p + \mu \nabla^2 \vec{v} \quad (5)$$

que ganhou esse nome graças aos trabalhos sobre movimento de fluidos realizados por C.L.M.H. Navier (1785-1836) e G.G Stokes (1819-1903); a Equação de condução de calor:

$$\frac{\partial T}{\partial t} + \bar{v} \cdot \bar{\nabla} T = k \nabla^2 T \quad (6)$$

e a Equação de conservação da massa:

$$\frac{\partial \rho}{\partial t} + \bar{\nabla}(\rho \bar{v}) = 0 \quad (7)$$

sendo T a temperatura, $p(T)$ a densidade (uma função da temperatura), \bar{v} a velocidade, \bar{F} a força gravitacional, p a pressão, e \bar{F}_c a força de Coriolis. Foi G.G. de Coriolis (1792-1843) quem introduziu as palavras: trabalho e energia cinética, com o sentido que tais possuem em Física. Em 1835, ele escreveu a segunda lei de Newton para um corpo em movimento, observando num referencial que gira com velocidade angular constante. A força de Coriolis é um termo que consta na equação que descreve o movimento desse corpo (MONTEIRO, 2006).

As grandezas que aparecem nas três equações anteriores são funções do tempo t e das coordenadas espaciais x , y e z . O eixo- x e o eixo- y formam um plano que representa a superfície da Terra; o eixo- z é perpendicular a esse plano.

As condições de contorno para uma camada atmosférica são:

$$\begin{aligned} T(x, y, 0, t) &= T_1 + \Delta T \\ T(x, y, h, t) &= T_1 \end{aligned} \quad (8)$$

de modo que a temperatura da camada de ar em contato com a superfície da Terra, em $z = 0$, é o maior por uma quantia $\Delta T > 0$ do que na altura $z = h$.

Uma simplificação usualmente feita é considerar que a dimensão horizontal do sistema é muito maior que do que sua dimensão vertical, de modo que os efeitos de borda são desprezíveis. Porém, ainda assim, a solução desse sistema de equações é bastante difícil.

Lorenz começou trabalhando com um sistema composto por doze equações diferenciais de primeira ordem, obtidas a partir das três equações anteriores. Demorou algum tempo até ele perceber que soluções estacionárias e periódicas eram inúteis para o seu propósito, pois o método linear é capaz de prever tais comportamentos. Para testar a capacidade de previsão do método linear, ele deveria encontrar uma solução aperiódica.

Em 1959, ele achou uma combinação de valores dos parâmetros que produzia tal solução. Quando aplicou o método linear para esses dados, ele verificou que a previsão estava longe de ser satisfatória, confirmando suas suspeitas iniciais (MONTEIRO, 2006). O passo de integração usado por Lorenz, em seus cálculos numéricos, era de “6 horas”. Normalmente, ele programava o computador para imprimir o valor das 12 variáveis a cada “dia” (ou seja, a cada quatro passos de integração). Para que todos esses números aparecessem numa única linha, ele imprimia apenas os três primeiros algarismos significativos de cada variável (MONTEIRO, 2006).

Com a intenção de analisar com mais detalhes um determinado cálculo, Lorenz resolveu refazê-lo, mas, para ganhar tempo, não desde o começo. Ele introduziu, no computador, uma linha de valores impressos que aparecia no meio da listagem obtida anteriormente. Surpreendentemente, depois de pouco tempo, os novos valores calculados não tinham relação alguma com aqueles

obtidos no cálculo precedente. Lorenz achou que seu computador estivesse com problemas e pensou em trocar uma válvula (MONTEIRO, 2006).

Antes de chamar alguém para realizar o conserto, ele voltou a examinar os dados e percebeu que, inicialmente, os novos valores coincidiam com os velhos, mas logo passavam a diferir na última casa, então na penúltima casa, depois na antepenúltima casa e assim por diante. Essa diferença aumentava rapidamente, até que, a partir de certo instante, as novas saídas não se pareciam mais com as antigas.

Passou-se algum tempo até Lorenz perceber que a impressora estava registrando os três primeiros algarismos dos resultados, mas o computador trabalhava, de fato, com seis algarismos. Lorenz compreendeu que havia, involuntariamente, introduzido um minúsculo erro inicial em razão do arredondamento, quando refez o cálculo. E esse erro crescia rapidamente à medida que o cálculo prosseguia, até chegar a um ponto que os valores calculados estavam radicalmente alterados, em comparação com os cálculos originais. Ele confirmou essa rápida divergência, realizando 40 simulações com pequenas diferenças nas condições iniciais (MONTEIRO, 2006).

Segundo Monteiro (2006), Lorenz resolveu o problema da imprevisibilidade atmosférica. Muito provavelmente, as equações completas, que descrevem com maior precisão a circulação atmosférica, apresentam essa mesma sensibilidade às condições iniciais, o que torna efetivamente impossível qualquer previsão do tempo em longo prazo.

O menor erro nas medidas das condições climáticas, num dado instante, pode comprometer a validade de qualquer previsão do tempo para os instantes seguintes. A dependência sensível das equações da circulação atmosféricas ficou conhecida como efeito borboleta. Segundo Lorenz, pequenas perturbações causadas pelo bater de asas de uma borboleta no Brasil podem provocar o surgimento de um tornado no Texas ou impedir o surgimento desse tornado.

Lorenz queria divulgar seus resultados, mas não estava satisfeito com seu sistema de doze equações. Um sistema mais simples seria mais interessante. Afinal, ele queria apenas mostrar a limitação do método linear de prever séries temporais complexas. Seu objetivo principal não era construir um modelo realista para a circulação atmosférica (MONTEIRO, 2006).

Em 1961, ele visitou Barry Saltzman, que trabalhava com um sistema de sete equações para o qual existia uma solução que parecia estar, eternamente, no regime transiente. Lorenz notou que os valores de quatro, das sete variáveis, logo se tornavam muito pequenos, e ficou curioso para saber se o sistema formado apenas pelas outras três variáveis também exibiria comportamento aperiódico. Saltzman autorizou o uso das equações por Lorenz.

Em seguida, Lorenz resolveu numericamente a versão simplificada com apenas três variáveis, e obteve também uma solução aperiódica, para os mesmos valores dos parâmetros escolhidos por Saltzman (MONTEIRO, 2006). O sistema de Lorenz é formado pelas Equações:

$$\begin{aligned}\frac{dX}{dt} &= f_1(X, Y, Z) = -\sigma X + \sigma Y \\ \frac{dY}{dt} &= f_2(X, Y, Z) = rX - Y - XY \\ \frac{dZ}{dt} &= f_3(X, Y, Z) = XY - bZ\end{aligned}\tag{9}$$

O parâmetro $\sigma \equiv \mu/k$ é conhecido como número de Prandtl, em homenagem a L. Prandtl (1875-1953) por seus trabalhos teóricos em aerodinâmica, $r \equiv R/R_c$ e $b \equiv 4/(1+a^2)$. Note que o

parâmetro r é proporcional a diferença de temperatura ΔT e b dependem das dimensões do sistema (MONTEIRO, 2006).

A variável $X(t)$ é proporcional à intensidade da convecção: $X = 0$ implica que não há movimento convectivo, ou seja, o calor é transportado apenas por condução; $X > 0$ implica circulação horária e $X < 0$, circulação anti-horária.

A variável $Y(t)$ é proporcional à diferença de temperatura entre as correntes ascendente e descendente do fluido (MONTEIRO, 2006).

A variável $Z(t)$ é proporcional à distorção do perfil vertical da temperatura, em relação ao perfil linear. Por exemplo, para $Z = 0$, a temperatura decresce linearmente, conforme se eleva pelo eixo- z , do seguinte modo:

$$T = T_1 + \Delta T \left(1 - \frac{z}{h} \right) \quad (10)$$

Saltzman escolhera $\sigma = 10$, que é um número característico de líquidos (para a água, $\sigma \cong 5$). Como meteorologista, Lorenz tomaria σ em torno de 1, o que é mais adequado para o ar. Entretanto, com esse valor de σ , Lorenz não encontraria a solução aperiódica almejada (MONTEIRO, 2006).

A Figura 2 exibe a solução do sistema de Lorenz para $\sigma = 10$, $r = 28$ e $b = 8/3$, a partir da condição inicial $X(0) = 1$, $Y(0) = 2$, $Z(0) = 3$. Essa solução está representada no espaço de fases, que é o espaço formado pelo eixo- X , pelo eixo- Y e pelo eixo- Z . A figura, cuja forma lembra as asas de uma borboleta, é o atrator desse sistema dinâmico.

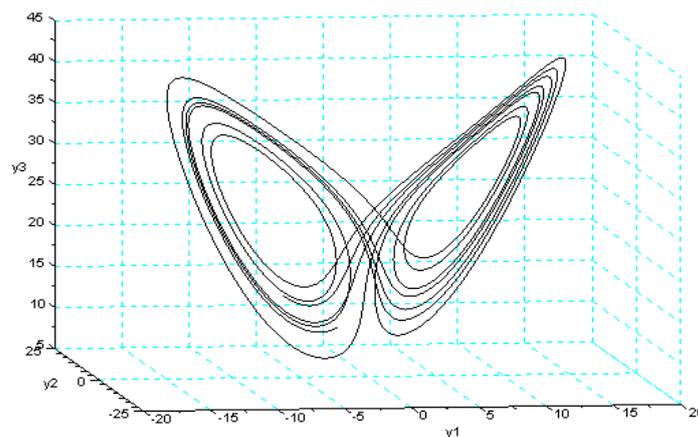


Figura 2 – O atrator do sistema de Lorenz.

Em 1963, foi publicado o artigo *Deterministic Nonperiodic Flow* de Lorenz, sobre um sistema de baixa dimensão que exibia comportamento caótico.

A invenção do computador, em meados do século XX marcou uma linha divisória no estudo de sistemas dinâmicos. Com o computador, pode-se realizar diversos experimentos com equações, calculando suas soluções a partir das condições iniciais e dos valores dos parâmetros desejados. A coleção de resultados numéricos obtidos ajuda a desenvolver uma intuição sobre os possíveis comportamentos que essas equações representam (MONTEIRO, 2006).

Segundo Monteiro (2006), não é surpreendente que sistemas complexos, como a dinâmica atmosférica, variação do preço de uma ação na bolsa de valores e a formação de trânsito urbano

exibam comportamentos irregulares, já que a evolução desses sistemas é influenciada por uma infinidade de fatores. Portanto, parece razoável que eles sejam, num certo sentido, imprevisíveis. Entretanto, comportamento caótico pode ocorrer em sistemas extremamente simples, como, por exemplo, na Equação de diferença:

$$x_{j+1} = \mu x_j (1 - x_j), \text{ com } x_j \in [0,1] \text{ e } j = 0,1,2,3,\dots \quad (11)$$

Essa equação, para $\mu = 4$, apresenta o mesmo tipo de sensibilidade às condições iniciais que o sistema de Lorenz. A descoberta de sistemas caóticos afetou o próprio método científico criado por Galileu. A abordagem tradicional para testar a validade de uma teoria consistia em verificar se as suas previsões se sustentavam quando comparadas com resultados experimentais. Todavia, se o sistema analisado é caótico, previsões de longo prazo é intrinsecamente impossível, o que dificulta a validação da teoria (MONTEIRO, 2006).

5. RESULTADOS E DISCUSSÃO

A solução do sistema de Lorenz para $\sigma = 10$, $r = 28$ e $b = 8/3$, a partir da condição inicial $X(0) = 1$, $Y(0) = 2$, $Z(0) = 3$. Essa solução é utilizada nos dois experimentos.

Os parâmetros utilizados na RNA MLP no experimento 1: número de neurônios de entrada igual a 15, número de camadas igual a 3, número de neurônios nas camadas ocultas igual a 100, taxa de aprendizagem constante igual a 0,1, fator de momento constante igual a 0, e número máximo de épocas igual a 50. O tempo de processamento foi de 6,4 segundos. A Figura 3 mostra a evolução temporal da variável $X(t)$ e a aproximação calculada pela RNA.

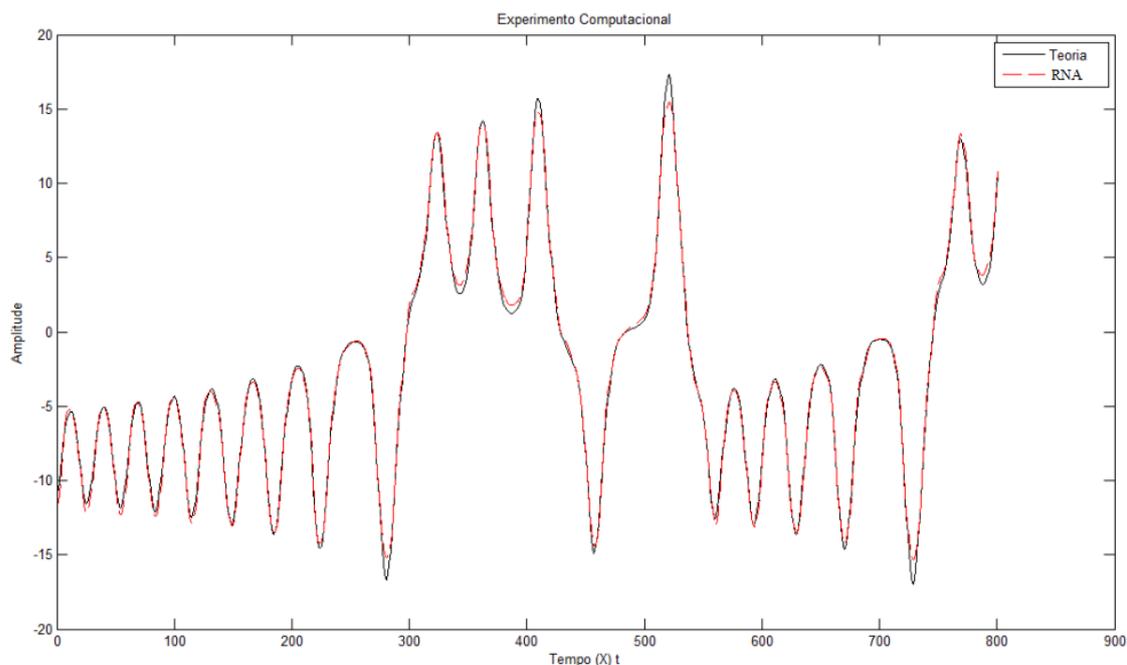


Figura 3 – Evolução temporal de $X(t)$ e a aproximação calculada pela RNA.

A Figura 4 mostra a curva de aprendizagem experimental (erro médio quadrado) com número máximo de 50 épocas.

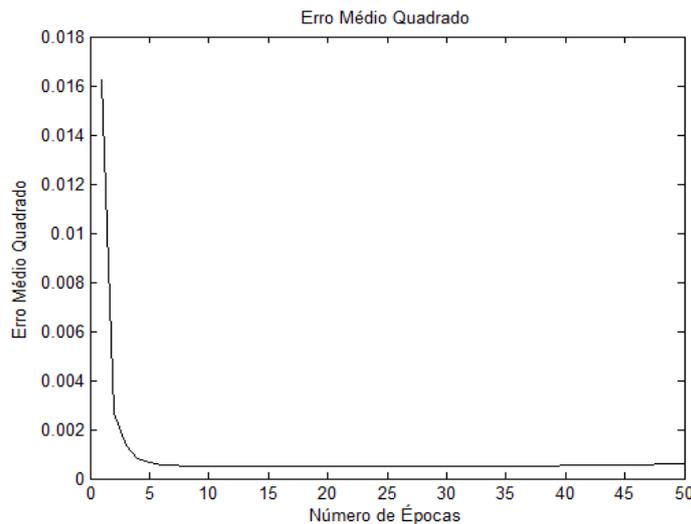


Figura 4 – Erro Médio Quadrado x 50 épocas.

Os parâmetros utilizados na RNA MLP no experimento 2: número de neurônios de entrada igual a 25, número de camadas igual a 3, número de neurônios nas camadas ocultas igual a 300, taxa de aprendizagem constante igual a 0,1, fator de momento constante igual a 0,5 e número máximo de épocas igual a 700. O tempo de processamento foi de 157,4 segundos. A Figura 5 mostra a evolução temporal da variável $X(t)$ e a aproximação calculada pela RNA no experimento 2.

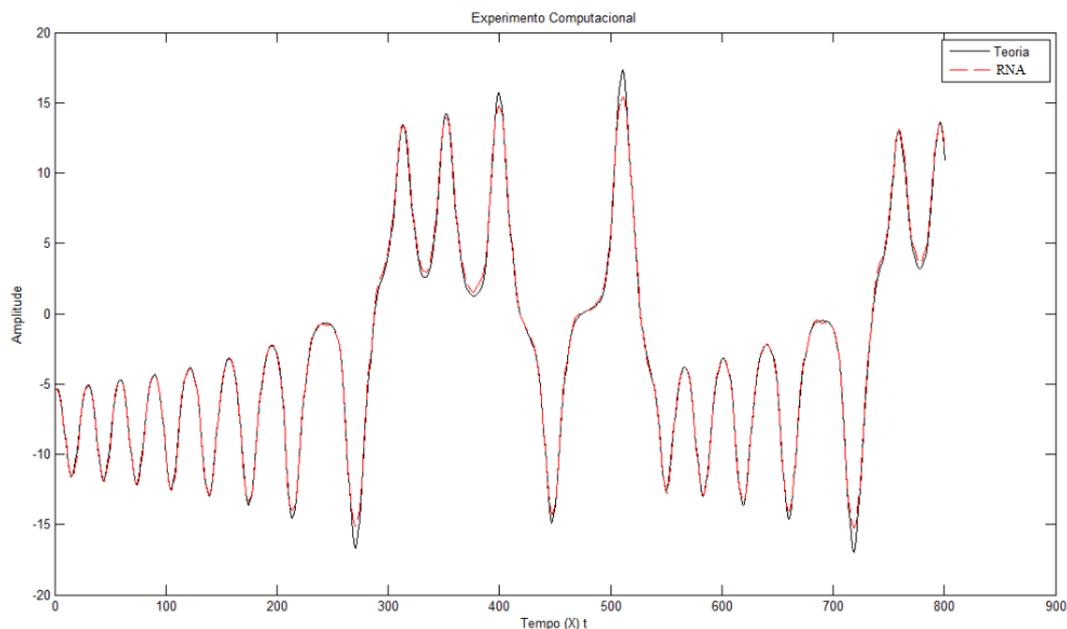


Figura 5 – Evolução temporal de $X(t)$ e a aproximação calculada pela RNA.

A Figura 6 mostra a curva de aprendizagem experimental (erro médio quadrado) com número máximo de 700 épocas.

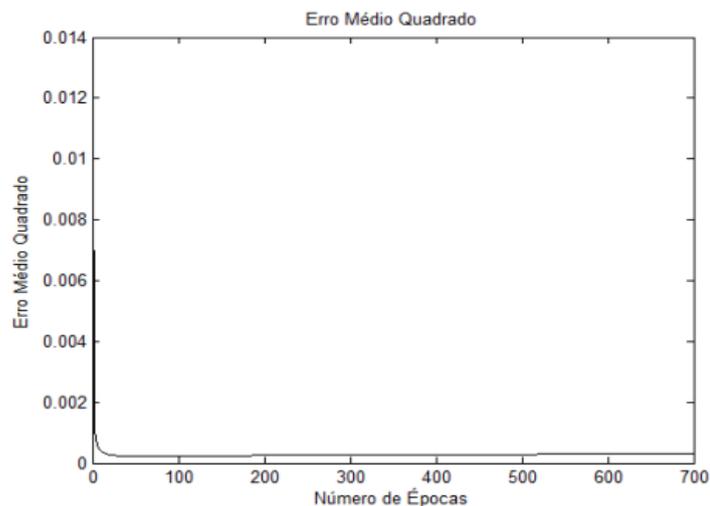


Figura 6 – Erro Médio Quadrado x 700 épocas.

4. CONCLUSÃO

Observa-se com base nos experimentos realizados com a Rede Neural Artificial, que os dois experimentos apresentaram boa aproximação da função de Evolução Temporal do Sistema de Lorentz. Pode-se concluir que as alterações nos parâmetros da RNA não apresentaram grande variação no resultado, exceto na variação de tempo de processamento.

Os erros médios quadrados nos dois experimentos estabilizaram próximo da trigésima época. As RNAs podem ser utilizadas para resolver numerosa quantidade de problemas reais complexos, entre esses: aproximação de funções, previsão e controle, reconhecimento de voz, reconhecimento óptico de caracteres, reconhecimento de padrões, mineração de dados, classificação etc.

O artigo não tem a intenção de esgotar o assunto, apenas contribuir com mais uma alternativa para resolução de problemas envolvendo aproximação de funções.

Espera-se que o artigo desperte interesse em outras áreas cujas pesquisas e aplicações são promissoras, tais como: engenharias, ciências exatas e da terra, biológicas, saúde, agrárias, sociais, humanas, linguística etc.

Como estudos futuros pretende-se realizar novos experimentos com outras funções de maior complexidade com o intuito de avaliar outros resultados experimentais. Pretende-se também utilizar outros tipos de RNAs para comparar o desempenho entre as diversas RNAs existentes.

6. REFERENCIAS

- AFFONSO, C. Aplicação de Redes Neuro Fuzzy ao Processamento de Polímeros na Indústria Automotiva. 2010. Dissertação (Mestrado) – Universidade Nove de Julho, Engenharia de Produção, São Paulo. 111 p.
- BIGUS, J. P. Data Mining with Neural Network: Solving Business Problems from Applications Development to Decision Support. Mcgraw-Hill, 1996.
- BRAGA, A. P.; CARVALHO A. C. P. L. F.; LUDERMIR, T. B. Redes Neurais Artificiais: Teoria e Aplicações. Rio de Janeiro. Ed. LTC, 2ª ed. Rio de Janeiro, 2011.

- FERREIRA, Ricardo Pinto. Combinação de Técnicas da Inteligência Artificial para Previsão do Comportamento do Tráfego Veicular Urbano na Cidade de São Paulo. 2011. Dissertação (Mestrado) – Universidade Nove de Julho, Engenharia de Produção, São Paulo. 107 p.
- GAGANIS, C. Classification techniques for the identification of falsified financial statements: a comparative analysis, *Intell. Syst. Account. Financ. Manag.* v.16, p. 207–229, 2009.
- HAYKIN, S. *Redes Neurais – Princípios e Práticas*. Bookman. 2a edição. Porto Alegre, 2001.
- HEBB, D. O. *The Organization of Behavior: Neuropsychological Theory*, N. Y.:Wiley, 1949.
- ITURRIAGA, F.J.L.; SANZ, I. P. Bankruptcy visualization and prediction using neural networks: a study of US commercial banks, *Exp. Syst. Appl.* 42, p. 2857–2869, 2015.
- KOVÁCS, Z. L. *Redes Neurais Artificiais - Fundamentos e Aplicações*. Editora Livraria da Física. 4ª edição. São Paulo, 2006.
- KRISTJANPOLLER, W.; MINUTOLO, M. C. Gold price volatility: a forecasting approach using the Artificial Neural Network–GARCH model, *Exp. Syst. Appl.* 2015.
- LABOISSIERE, L. A.; FERNANDES, R. A.; LAGE, G. G. Maximum and minimum stock price forecasting of Brazilian power distribution companies based on artificial neural networks, *Appl. Soft Comput.* 2015.
- LARRAIN, J. M.; TURNER, C. The treasury bill rate, the great recession, and neural networks estimates of real business sales, *Proc. Comput. Sci.* 36, p. 227–233, 2014.
- LORENTZ, E. N. Deterministic nonperiodic flow, *J. Atmosph. Sci.*, v.20. p. 130-141, 1963.
- MARTINIANO, A. S.; FERREIRA, R. P.; AFFONSO, C.; SASSI, R. J. Aplicação de uma Rede Neuro Fuzzy na Previsão do Absenteísmo no Trabalho. CISTI'2012 - 7ª Conferencia Ibérica de Sistemas y Tecnologías de Información, 2012.
- MATLAB ® version 7.8.0 R2009a, 2009.
- MCCULLOCH, W. S.; PITTS, W. H. A logical calculus of the ideas immanent in nervous activity, *Bulletin of mathematical biophysics*, v. 5, p. 115-133, 1943.
- MENDEL, J. M.; MCLAREN, R. W. Reinforcement-learning control and pattern recognition systems. *Adaptive, Learning and Pattern Recognition Systems*. Academic Press, New York e London, Cap. 8, p. 287–318, 1970.
- MONTEIRO, L. H. A. *Sistemas Dinâmicos*. 2. ed. São Paulo: Editora Livraria da Física, 2006.
- RATH, S.; JAGADEV, A. K.; NAYAK, M. R. Performance analysis of stock market using artificial neural network, *Int. J. Appl. Eng. Res.* 10, 2015.
- ROSENBLATT, M. The Perceptron: A probabilistic model for information storage and organization in the Brain. *Psychological review*, v.65, n.6, p. 386-408, 1958.
- RUSSEL, S.; NORVIG, P. *Inteligência Artificial*. Campus. 3. ed. São Paulo, 2013.
- SASSI, Renato José. Uma Arquitetura Híbrida para Descoberta de Conhecimento em Bases de Dados: Teoria dos Rough Sets e Redes Neurais Artificiais Mapas Auto-Organizáveis. 169 f. Tese (Doutorado) – Escola Politécnica da Universidade de São Paulo, Departamento de Engenharia Elétrica, São Paulo, 2006.
- SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A. *Redes Neurais Artificiais para Engenharia e Ciências Aplicadas*. SP: Artliber, 2010.
- SIMÕES, M. G.; SHAW, I. S. *Controle e Modelagem fuzzy*. São Paulo: Blucher: FAPESP, 2007.
- TEORIA DO CAOS. Disponível em: < <http://www.brasilecola.com/fisica/teoria-caos.htm#> > Acesso em 04 fev. 2016.