

Proposta de uma Classe de Perceptrons Híbridos com Aprendizagem baseada em Gradiente Descendente

Ricardo de A. Araújo^{1,2} (Autor)

Silvio Meira¹ (Orientador), Adriano L. I. Oliveira¹ (Co-orientador)

¹CIn, Universidade Federal de Pernambuco, Recife, PE, Brasil

²DEINFO, Instituto Federal do Sertão Pernambucano, Ouricuri, PE, Brasil

Resumo. *Este trabalho apresenta um modelo, chamado de dilation-erosion-linear perceptron (DELP), que consiste de uma combinação linear entre operadores não-lineares e um operador linear, para superar o dilema do passeio aleatório no problema de previsão de séries temporais financeiras. Para projetar o DELP foi apresentado um método baseado em gradiente descendente utilizando ideias do algoritmo de retropropagação do erro e uma abordagem sistemática para superar o problema da não-diferenciabilidade das operações não-lineares empregadas no modelo. Também, foi incluída uma etapa adicional para ajustar distorções de fase temporais que ocorrem na reconstrução do espaço de fase de fenômenos temporais financeiros. Por fim, uma análise experimental foi conduzida utilizando um conjunto relevante de séries financeiras, onde foram utilizadas cinco métricas e uma função de avaliação para mensurar o desempenho preditivo do modelo, e os resultados alcançados superaram aqueles obtidos por técnicas consolidadas na literatura.*

Abstract. *This work presents a model, called the dilation-erosion-linear perceptron (DELP), which consists of a linear combination of nonlinear and linear operators, to overcome the random walk dilemma (RWD) in the financial time series prediction problem. In order to design the proposed DELP, we present a gradient steepest descent method using the back propagation (BP) algorithm ideas and a systematic approach to overcome the problem of nondifferentiability of nonlinear operators employed within model. Also, we have included an additional step to adjust time phase distortions that occur in the phase space reconstruction of financial temporal phenomena. At the end, an experimental analysis was conducted using a relevant set of financial time series, where we used five relevant metrics and an evaluation function to assess the prediction performance, and the achieved results overcame those obtained by established techniques in the literature.*

1. Resumo da Dissertação

Neste trabalho, o problema de previsão de séries temporais foi abordado com o estudo de várias técnicas inspiradas na estatística e inteligência computacional (*computational intelligence*, CI). Os modelos estatísticos, lineares e não-lineares, são os mais popularmente encontrados na literatura de previsão de séries temporais. Neste contexto, destacam-se os modelos autoregressivos integrados de médias móveis (*autoregressive integrated moving average*, ARIMA), que são os mais difundidos e utilizados em aplicações de previsão financeira.

Entretanto, devido a linearidade dos modelos ARIMA é possível identificar uma limitação prática de sua utilização, uma vez que os fenômenos temporais financeiros são comumente não-lineares. Analisando os modelos estatísticos não-lineares, pode-se verificar uma limitação prática de sua utilização, uma vez que a não-linearidade de tais modelos não implica em um ganho significativo, em termos de desempenho preditivo, quando comparado ao desempenho de modelos estatísticos lineares. Neste contexto, uma alternativa promissora surgiu a partir do desenvolvimento de modelos baseados em CI para previsão de séries temporais.

Dentre estes modelos, foram analisados aqueles mais comumente encontrados em aplicações de séries financeiras, bem como os modelos que na literatura obtiveram os melhores resultados preditivos, que foram os modelos baseados em redes neurais do tipo *perceptron* multicamadas (*multilayer perceptron*, MLP). No entanto, as redes MLP requerem a definição de um conjunto de parâmetros bastante difícil de se determinar. Além disso, no caso particular do problema de previsão séries temporais financeiras utilizando redes MLP, um elemento crucial para previsão é a definição dos retardos temporais para representar a série.

A partir deste ponto, foram investigados o algoritmo genético modificado (*modified genetic algorithm*, MGA) e o otimizador de enxame de partículas (*particle swarm optimizer*, PSO) para configuração e ajuste automático de todos os parâmetros das redes MLP, gerando sistemas híbridos (MGA-MLP e PSO-MLP). No entanto, como o espaço de busca dos parâmetros das redes MLP é bastante complexo, ambas as técnicas podem não ser capazes de ajustar e treinar, simultaneamente e eficientemente, este tipo particular de rede neural. Portanto, foram investigados sistemas híbridos otimizados, o método *time-delay added evolutionary forecasting* (TAEF) e o método *swarm-based hybrid intelligent forecasting* (SHIF), que utilizam o MGA e o PSO, respectivamente, para ajuste das redes MLP, e utilizam técnicas baseadas em gradiente descendente para treiná-las. Estes modelos híbridos otimizados apresentaram resultados promissores e atualmente são consideradas as melhores técnicas de previsão encontradas na literatura.

Entretanto, mesmo com as técnicas sofisticadas e metodologias híbridas que têm sido recentemente propostas na literatura, uma limitação surge de todos estes modelos de previsão no caso particular de séries temporais financeiras, conhecido como dilema do passeio aleatório. Este dilema ocorre quando se tenta prever qualquer série temporal financeira utilizando um modelo arbitrário de previsão, onde conseqüentemente as previsões geradas apresentam um característico atraso de um passo a frente em relação aos valores reais da série temporal, isto é, possuem uma distorção de fase temporal na reconstrução do espaço de fase de um determinado fenômeno temporal financeiro.

Devido a este fato, alguns pesquisadores argumentaram que séries temporais financeiras são geradas por processos aleatórios e não podem ser previstas. Entretanto, a justificativa para tal afirmação se baseia no fato que nenhum modelo de previsão proposto na literatura é capaz de prever estas séries sem o atraso de um passo da previsão em relação aos valores reais da série. Portanto, utilizou-se um conjunto relevante de séries temporais financeiras comumente utilizadas na literatura na tentativa de investigar melhor este tipo de série. A análise dos *lagplots* destas séries nos fornece a demonstração prática que séries temporais financeiras são geradas por componentes não-lineares complexas subdominantes e uma componente linear dominante. Este argumento nos fornece as indicações necessárias para defender a não aleatoriedade destas séries.

Nesta análise realizada, verifica-se que as séries financeiras investigadas não são geradas por processos aleatórios, uma vez que foi encontrada a presença de relacionamentos não-lineares subdominantes complexos embutidos em relacionamentos lineares dominantes. Se tais séries fossem unicamente aleatórias, não seria possível identificar quaisquer relacionamentos não-lineares em retardos temporais de alta ordem, isto é, seria possível identificar apenas um aumento na variância nos pontos destes gráficos, uma vez que em um modelo de passeio aleatório a variância aumenta linearmente com o aumento da ordem dos retardos temporais.

Assim, pode-se concluir que séries financeiras têm um comportamento bastante similar a séries de passeio aleatório em retardos de baixa ordem. No entanto, com o aumento da ordem dos retardos temporais, tem-se claras evidências de relacionamentos não-lineares complexos presentes nestas séries. Neste sentido, pode-se notar que este tipo de série não é gerado por um processo aleatório, mas por uma combinação de processos lineares dominantes e processos não-lineares subdominantes. Note que o relacionamento linear dominante é sensível aos eventos em um passado bastante recente, o que leva a um comportamento similar a um processo aleatório devido a fortes inferências especulativas no mercado financeiro. Também, é possível verificar que o relacionamento

não-linear subdominante é responsável por eventos de médio/longo prazo, o que leva a uma percepção correta do fenômeno temporal sem a inferência especulativa (aqui pode-se observar o comportamento de médio/longo prazo da série, sendo este responsável pelo aparecimento destas estruturas não-lineares bem definidas nos *lagplots*). Também, outra observação interessante apresentada nesta análise é que tais componentes podem ser eficientemente aproximadas em termos de uma combinação balanceada entre operadores lineares de resposta finita ao impulso (componente linear) e operadores não-lineares crescentes (componente não-linear). A justificativa da estimativa da componente não-linear utilizando operadores crescentes é devido ao fato que a análise destas séries enfatizando o critério de médio/longo prazo revelou que o comportamento do mercado financeiro é naturalmente crescente. Por fim, foi apresentado um conjunto de métricas com características diversas e uma função de avaliação para medir o desempenho preditivo global dos modelos investigados.

Definidas as principais motivações deste trabalho, foi proposta uma classe de *perceptrons* híbridos baseado nos princípios da morfologia matemática (*mathematical morphology*, MM) no contexto de teoria de reticulados (*lattice theory*) para superar o dilema do passeio aleatório (*random walk dilemma*, RWD) no problema de previsão de séries temporais financeiras. A MM sobre o contexto de teoria de reticulados foi escolhida para fundamentar o modelo proposto devido à sua natureza inerentemente crescente. Vale mencionar que a justificativa teórica da possibilidade de construção do futuro se fundamenta no trabalho de Takens, que mostrou, matematicamente, que é possível reproduzir qualquer fenômeno temporal não aleatório no tempo futuro a partir de suas observações no tempo passado, isto é, a partir da escolha de uma certa dimensionalidade mínima necessária para uma correta caracterização do sistema gerador deste fenômeno.

O modelo proposto foi chamado de *perceptron* de dilatação-erosão-linear (*dilation-erosion-linear perceptron*, DELP), uma vez que este consiste de uma combinação linear entre operadores não-lineares (do tipo morfológicos no contexto de teoria de reticulados) e um operador linear (do tipo resposta finita ao impulso). A ideia de propor este modelo é fundamentada pela análise das séries temporais financeiras realizada. Assim, o DELP foi projetado de forma a conter toda informação necessária para se prever séries temporais financeiras, uma vez que sabe-se a priori os tipos de relacionamento que governam esse tipo particular de séries temporais (relacionamento linear dominante e um relacionamento não-linear subdominante com comportamento crescente).

Para projetar o DELP (processo de aprendizagem), foi proposto um método de gradiente descendente utilizando ideias do algoritmo de retropropagação do erro (*back propagation*, BP) e uma abordagem sistemática para superar o problema da não-diferenciabilidade das operações morfológicas de dilatação e erosão, uma vez que estes operadores não são diferenciáveis de forma usual. Também, no processo de aprendizagem do DELP, foi incluída uma etapa adicional, definida como procedimento de correção de fase (*phase fix procedure*, PFP), para ajustar distorções de fase temporais que ocorrem na reconstrução do espaço de fase de fenômenos financeiros.

Para se estabelecer um nível de referência para o desempenho preditivo do modelo proposto, foram inicialmente realizados experimentos com os modelos ARIMA, haja vista que estes são modelos lineares comumente utilizados em aplicações no mercado financeiro. Posteriormente, foram realizados vários experimentos com redes MLPs, onde os ajustes e configurações aplicados foram escolhidos através de exaustivos experimentos preliminares com o objetivo de otimizar o modelo. Depois, foram realizados os experimentos com os sistemas híbridos MGA-MLP e PSO-MLP e os sistemas híbridos otimizados TAEF e SHIF. Por fim, foram realizados os experimentos com o modelo proposto.

Analisando os resultados alcançados, pode-se verificar que as redes MLP obtiveram um desempenho preditivo superior aos modelos ARIMA. No entanto, mesmo aplicando o procedimento de correção de fase (que ajusta distorções de fase temporais na representação das séries financeiras), ambos os modelos seguem o mesmo comportamento de um modelo de passeio aleatório. Observando este comportamento, pode-se

assumir que este resultado é um reflexo da escolha incorreta dos parâmetros do modelo de previsão (sendo determinados a partir de procedimentos manuais e não otimizados).

Quando se observa os resultados obtidos pelos sistemas híbridos MGA-MLP e PSO-MLP, verifica-se que estes obtiveram desempenho levemente superior aos modelos MLP e ARIMA. No entanto, mais uma vez, mesmo aplicando o procedimento de correção de fase, ambos os modelos também seguem o mesmo comportamento de um modelo de passeio aleatório. Este comportamento vem provavelmente do fato do espaço de busca para os parâmetros da rede MLP constitui um universo extremamente vasto e de dimensionalidade alta, necessitando de uma grande quantidade de tempo para um correto treinamento e ajuste, simultâneo e automático, da rede MLP. Mesmo assim, estes modelos híbridos têm a vantagem de não necessitarem da interação de um especialista humano para construção do modelo final de previsão.

Uma explicação plausível para a incapacidade de ajuste das distorções de fase temporais destes modelos de previsão é que o procedimento de correção de fase depende da complexidade da informação contida na série temporal e a habilidade de definir precisamente os melhores parâmetros do modelo de previsão para estimar os valores reais da série, ou seja, o sucesso do procedimento de correção de fase é fortemente dependente do ajuste acurado dos parâmetros do modelo de previsão, bem como do próprio modelo utilizado para o propósito de previsão.

Ao observar os resultados alcançados com os modelos híbridos otimizados (TAEF e SHIF), verificou-se um desempenho preditivo superior aos modelos consolidados na literatura, onde o procedimento de correção de fase se mostrou capacitado para o ajuste da distorção de fase temporal, o que evidencia que alguns modelos de previsão só conseguem superar o problema quando são corretamente ajustados com técnicas de busca globais combinadas com técnicas de busca locais para refinamento dos parâmetros globais.

Observando os resultados obtidos nos experimentos realizados, pode-se notar que o modelo proposto conseguiu superar, em termos de qualidade de previsão, os resultados alcançados pelos modelos consolidados na literatura de previsão de séries financeiras, considerando todas as métricas, individualmente, bem como considerando a função de avaliação global. Pode-se notar que, devido a natureza linear e não-linear do DELP (característica inerente a séries financeiras), e a inclusão do procedimento de correção de fase do algoritmo de treinamento do modelo proposto, este foi capaz de ajustar mais precisamente a distorção de fase temporal de séries temporais financeiras em relação ao ajuste alcançado pelos modelos TAEF e SHIF. Em outras palavras, o modelo proposto foi capaz de superar, mais eficientemente, o dilema do passeio aleatório para previsão de séries temporais financeiras.

Uma explicação mais detalhada de tal comportamento é que o modelo proposto foi construído com características similares as que regem o fenômeno gerador de séries temporais financeiras, o que reduz o espaço de busca por um modelo ótimo de previsão. Tal fato associado a inclusão do procedimento de correção de fase no algoritmo de treinamento do DELP leva a um ajuste mais preciso das distorções de fase temporais, uma vez que a complexidade da informação contida nas séries financeiras é mapeada diretamente nas equações do DELP.

Vale mencionar que, em todas as séries temporais analisadas, verifica-se que o modelo proposto utilizou em torno de 98% da componente linear e em torno de 2% da componente não-linear. Como o modelo proposto é capaz de selecionar o percentual de utilização de cada operador em particular, pode-se verificar que o mesmo estimou tais séries a partir de uma combinação balanceada de operadores lineares dominantes e operadores não-lineares subdominantes. Tal fato confirma a análise dos *lagplots* que séries financeiras são geradas por uma combinação de relacionamentos lineares dominantes com relacionamentos não-lineares subdominantes, uma vez que o modelo proposto é visto como um mapeamento direto dos relacionamentos contidos em séries temporais financeiras.

Também, pode-se observar que o modelo DELP não teve nenhuma similaridade

com um modelo de passeio aleatório, uma vez que em nenhuma série temporal analisada ocorreu o característico atraso de um passo em relação aos valores reais da série. Também, vale mencionar que o primeiro retardo temporal não foi selecionado para prever qualquer série com o modelo proposto, uma vez que uma estrutura de passeio aleatória é necessária para se utilizar o procedimento de correção de fase, pois a chave desta etapa é a previsão de dois passos para o ajuste do atraso de um passo. Note que, uma das principais vantagens do modelo proposto, além de seu desempenho preditivo (sem a utilização de técnicas evolucionárias e/ou híbridas em seu processo de aprendizagem) bem mais acurado quando comparado as melhores técnicas de previsão reportadas na literatura, é sua estrutura computacional bastante simples e a sua habilidade de estimar, de forma balanceada, as componentes linear e não-linear, reproduzindo fielmente o fenômeno gerador de séries financeiras. Desta forma, o DELP é modelo viável, em termos práticos e científicos, para solucionar o dilema do passeio aleatório para previsão financeira.

Por fim, seguem as principais contribuições deste trabalho: **i)** Investigação de modelos estatísticos lineares e não-lineares, bem como modelos baseados em inteligência artificial, aplicados para previsão de séries temporais financeiras, **ii)** Identificação, a partir da análise dos *lagplots*, que fenômenos temporais financeiros são gerados por uma combinação de componentes lineares dominantes e componentes não-lineares complexas sub-dominantes, sendo esta a justificativa da possibilidade de se prever eficientemente séries financeiras, **iii)** Proposta de uma classe de *perceptrons* híbridos baseado nos princípios da MM no contexto de teoria de reticulados para superar o dilema do passeio aleatório no problema de previsão de séries temporais financeiras, **iv)** Demonstração da possibilidade de se estimar a componente não-linear de séries temporais financeiras utilizando os operadores morfológicos crescentes, **v)** Definição formal do *perceptron* de dilatação-erosão-linear (*dilation-erosion-linear perceptron*, DELP), **vi)** Desenvolvimento de uma abordagem sistemática para superar o problema da não-diferenciabilidade das operações morfológicas que serão utilizadas no algoritmo de treinamento do DELP, **vii)** Desenvolvimento do processo de aprendizagem do DELP utilizando o método de gradiente descendente e ideias do algoritmo de retropropagação do erro, **viii)** Inclusão do procedimento de correção de fase no processo de aprendizagem do DELP visando o ajuste de distorções de fase temporais que ocorrem na reconstrução do espaço de fase de fenômenos temporais financeiros (comportamento caracterizado por um atraso de um passo a frente da previsão em relação aos valores reais da série – “dilema do passeio aleatório”), **ix)** Realização de uma análise experimental extensiva com o modelo proposto e outros modelos apresentados na literatura, e **x)** Investigação do desempenho do modelo proposto utilizando um conjunto relevante de séries financeiras a fim de demonstrar o comportamento do modelo proposto como solução viável para o dilema do passeio aleatório.

Como produto deste trabalho foram publicados os artigos em periódicos: [de A. Araújo 2012c], [de A. Araújo 2012a], [de A. Araújo 2012b], [de A. Araújo et al. 2012b], [de A. Araújo et al. a], [de A. Araújo], [de A. Araújo et al. b], [de A. Araújo 2011a], [de A. Araújo et al. 2011i] e [de A. Araújo 2011b], bem como foram publicados os seguintes artigos em congressos e simpósios: [de A. Araújo et al. 2012a], [de A. Araújo et al. 2012c], [de A. Araújo et al. 2012d] (Artigo premiado entre os melhores da conferência), [de A. Araújo et al. 2011g], [de A. Araújo et al. 2011b], [de A. Araújo et al. 2011a], [de A. Araújo et al. 2011d] (Artigo premiado entre os melhores da conferência), [de A. Araújo et al. 2011e], [de A. Araújo et al. 2011c], [de A. Araújo et al. 2011h] e [de A. Araújo et al. 2011f].

Referências

- de A. Araújo, R. Evolutionary learning processes to design the dilation-erosion perceptron for weather forecasting. *Neural Processing Letters – Accepted For Publication*.
- de A. Araújo, R. (2011a). A class of hybrid morphological perceptrons with application in time series forecasting. *Knowledge-Based Systems*, 24(4):513–529.
- de A. Araújo, R. (2011b). Translation invariant morphological time-lag added evolutionary forecasting method for stock market prediction. *Expert Systems with Applications*, 38(3):2835–2848.

- de A. Araújo, R. (2012a). Hybrid morphological methodology for software development cost estimation. *Expert Systems with Applications*, 39:6129–6139.
- de A. Araújo, R. (2012b). A morphological perceptron with gradient-based learning for brazilian stock market forecasting. *Neural Networks*, 28:61–81.
- de A. Araújo, R. (2012c). A robust automatic phase-adjustment method for financial forecasting. *Knowledge-Based Systems*, 27:245–261.
- de A. Araújo, R., de Oliveira, A. L. I., Soares, S. C. B., and de L. Meira, S. R. (2011a). Designing dilation-erosion perceptrons with differential evolutionary learning for air pressure forecasting. In *International Joint Conference on Neural Networks*. IEEE.
- de A. Araújo, R., de Oliveira, A. L. I., Soares, S. C. B., and de L. Meira, S. R. (2011b). Dilation-erosion perceptrons with evolutionary learning for weather forecasting. In *IEEE International Conference on Systems, Man, and Cybernetics*. IEEE.
- de A. Araújo, R., de Oliveira, A. L. I., Soares, S. C. B., and de L. Meira, S. R. (2011c). An evolutionary approach to design dilation-erosion perceptrons for stock market indices forecasting. In *ACM Genetic and Evolutionary Computation Conference*. ACM.
- de A. Araújo, R., de Oliveira, A. L. I., Soares, S. C. B., and de L. Meira, S. R. (2011d). Gradient-based morphological approach for software development cost estimation. In *International Joint Conference on Neural Networks*. IEEE.
- de A. Araújo, R., de Oliveira, A. L. I., Soares, S. C. B., and de L. Meira, S. R. (2011e). A method to overcome the random walk dilemma in financial forecasting. In *International Symposium on Forecasting*.
- de A. Araújo, R., de Oliveira, A. L. I., Soares, S. C. B., and de L. Meira, S. R. (2011f). Quantum-inspired method for financial time series forecasting. In *IEEE Congress on Evolutionary Computation*. IEEE.
- de A. Araújo, R., de Oliveira, A. L. I., Soares, S. C. B., and de L. Meira, S. R. (2011g). A quantum-inspired morphological approach to solve the random walk dilemma for financial forecasting. In *Congresso Brasileiro de Inteligência Computacional*.
- de A. Araújo, R., de Oliveira, A. L. I., Soares, S. C. B., and de L. Meira, S. R. (2011h). Swarm-based evolutionary learning for dilation-erosion perceptrons with application in natural phenomena prediction. In *IEEE Congress on Evolutionary Computation*. IEEE.
- de A. Araújo, R., Oliveira, A. L. I., and Meira, S. A quantum-inspired evolutionary learning process to design dilation-erosion perceptrons for financial forecasting. *Learning and Nonlinear Models – Accepted For Publication*.
- de A. Araújo, R., Oliveira, A. L. I., and Meira, S. A swarm-based evolutionary morphological approach for binary classification problems. *Learning and Nonlinear Models – Accepted For Publication*.
- de A. Araújo, R., Oliveira, A. L. I., and Meira, S. (2012a). A dilation-erosion-linear perceptron for bovespa index prediction. *International Conference on Intelligent Data Engineering and Automated Learning*.
- de A. Araújo, R., Oliveira, A. L. I., and Meira, S. (2012b). An evolutionary morphological approach for software development cost estimation. *Neural Networks*, 32:285–291.
- de A. Araújo, R., Oliveira, A. L. I., and Meira, S. (2012c). A hybrid model for s&p500 index forecasting. *International Conference on Artificial Neural Networks*.
- de A. Araújo, R., Oliveira, A. L. I., and Meira, S. (2012d). A hybrid neuron with gradient-based learning for binary classification problems. *Encontro Nacional de Inteligência Artificial - ENIA*.
- de A. Araújo, R., Oliveira, A. L. I., and Soares, S. C. B. (2011i). A shift-invariant morphological system for software development cost estimation. *Expert Systems with Applications*, 38(4):4162–4168.