

APRENDIZAGEM AUTÔNOMA PARA GERENCIAMENTO DE UMA BOLSA DE VALORES SIMPLIFICADA

LEILTON S. LEMOS, JEAN C. SOUZA, MARCOS RAW, CARLOS H. C. RIBEIRO E KARL H. KIENITZ

Departamento de Engenharia Elétrica e Computação

Instituto Tecnológico De Aeronáutica - ITA.

Praça Mal. Eduardo Gomes, 50, Vila das Acácias, São José dos Campos, SP, BR, 12.228-900

E-mails: leiltonz@zipmail.com.br, jean_souza@hotmail.com, marcosraw@bol.com.br, carlos@comp.ita.br, kienitz@ele.ita.br

Resumo— Atualmente, diversos estudos estão sendo desenvolvidos visando a predição da movimentação do mercado de ações utilizando Redes Neurais Artificiais com uma arquitetura complexa e um grande número de entradas de dados, conseguindo-se resultados próximos aos de um analista de mercado. Neste trabalho, é apresentada uma nova proposta de uma Rede Neural com Aprendizado por Reforço (RNAF) usando uma estrutura simples, constituída de apenas três neurônios, nenhuma camada escondida e um número reduzido de entradas de dados. Os resultados obtidos mostram a eficiência de uma rede com atualização de sinapses usando *Aprendizado por Reforço* neste tipo de aplicação. Devido a sua característica de *aprendizado temporal*, esta rede mostrou-se mais flexível a alterações de dinâmica no sistema. Foram feitas comparações com os resultados obtidos com uma Rede Perceptron com mesma arquitetura e, ainda, com dados de empresas corretoras de valores, que realizam operações financeiras no mercado mundial. Evidenciou-se a eficiência para a análise de sistemas com dinâmica temporal, bem como a redução do conjunto de entradas de dados devido a esta característica.

Abstract— At present, several studies are being developed seeking the stock market prediction using Artificial Neural Network of complex architecture and a large number of data entries. In this work, a Reinforcement Learning based Neural Network (*RLNN*) was used with a simple structure, containing only three neurons, no hidden layer and a low number of data entries. The results obtained show the efficiency of the net in this kind of application. Due to its temporal learning ability, this net is more flexible to dynamic changes in the system. The results were compared with that obtained in a Perceptron Network with the same architecture of the RLNN and also, compared with data from stock-brokers companies that operate in the world market. The efficiency for analysis of systems with temporal dynamics was evidenced, as well as data entry reduction due to this characteristic.

Keywords— neural networks, reinforcement learning, autonomous learning, stock market

1 Introdução

Vários estudos têm sido desenvolvidos para a predição do movimento do mercado de ativos, grande parte deles utilizando redes neurais para efetuar tal predição. No entanto, tais abordagens exigem um grande número de entradas de dados e uma arquitetura complexa. Buscou-se portanto utilizar um número menor de dados de entrada, baseando-se na teoria de DOW (Blum, 1992) largamente utilizada na análise gráfica tipo DAYTRADE. Foram realizados testes envolvendo 54 ativos negociados na BOVESPA, escolhidos por terem histórico de negociação completo ou próximo disso (foram analisados no total 699 pregões de 02/01/1998 à 27/10/2000 aproximadamente 37750 fechamentos). Escolheu-se como agente previsor, uma Rede Neural com Aprendizado por Reforço. Comparou-se esta com uma Rede Perceptron com a mesma arquitetura de rede e, também, com os índices de grandes investidores.

2 A Bolsa de Valores

Cabem aqui, algumas prévias definições para um melhor entendimento dos termos usados no Mercado de Capitais:

Títulos – São ativos de renda variável emitidos por sociedades anônimas, que representam a menor fração do capital da empresa emitente. O investidor em ativos é co-proprietário da sociedade anônima da qual é acionista, participando dos seus resultados. Os ativos são conversíveis em dinheiro, a qualquer tempo, pela negociação em bolsas de valores.

Carteira de Ativos – Conjunto de ativos de diferentes empresas, de propriedade de pessoas físicas ou jurídicas.

Ativo em Alta – Quando o valor de fechamento atual é superior ao valor de fechamento anterior.

Ativo em Baixa – Quando o valor de fechamento atual é inferior ao valor de fechamento anterior.

Bolsa de Valores – Associação sem fins lucrativos que mantém um sistema adequado à realização de transações de ativos.

BOVESPA – Bolsa de Valores de São Paulo.

IBOVESPA – Índice da Bolsa de Valores de São Paulo, que mede a lucratividade de uma carteira teórica de ativos.

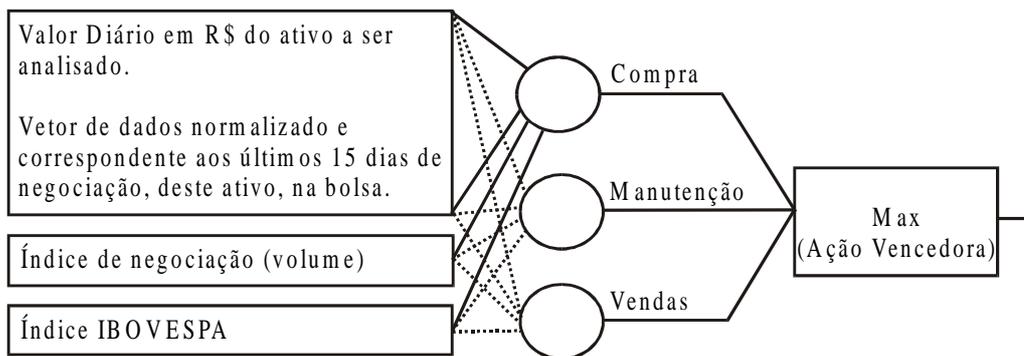


figura 1 - Arquitetura da Rede

3 A Rede Neural

O layout da rede (figura 2) foi definido como uma rede com 3 neurônios de entrada, onde a saída de cada um deles corresponde à utilidade de uma das três ações possíveis : Compra, Venda ou Manutenção. Considera-se a ação vencedora aquela que obtiver a maior utilidade como resposta. Há de se salientar que, não há camadas escondidas e os neurônios de saída são também os de entrada.

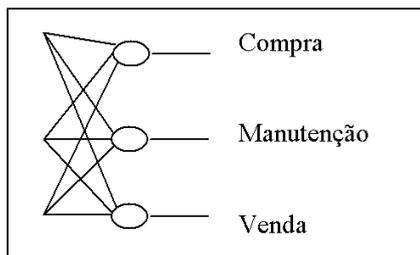


figura 2 - Layout da rede

3.1 Entradas da Rede Neural

Vale enfatizar que as entradas da rede não correspondem a estados (usual em Aprendizado por Reforço) e sim, a valores ou índices dos ativos. Os dados foram classificados em três categorias : Valor Diário, Índice de Negociação e IBOVESPA.

1.1.1 Valor Diário

Com base na teoria de DOW (Blum,) arbitrou-se pelo uso dos últimos 15 dias de negociação pois, na técnica de análise gráfica, os intervalos de análise variam de 8 a 13 dias. O valor diário foi escolhido como sendo o valor de fechamento do dia em análise e, para evitar a saturação das sinapses devido à variação destes valores, eles foram normalizados antes de inseridos na rede. Devido às flutuações diárias destes valores e à impossibilidade de obtê-los em tempo real, a rede executa somente uma escolha de ação diária para cada ativo, ocorrendo esta após o fechamento do pregão do dia anterior . Para inserir os dias na rede foi usado um vetor de dados que contém os valores do dia atual até o 14º dia anterior.

1.1.2 Índice de Negociação

O volume negociado corresponde ao número total de ativos negociados no dia (seja por compra ou venda). O índice de negociação é resultado de uma comparação entre o volume negociado no dia e a média dos 5 dias anteriores. Devido a grande diferença na quantidade disponível para cada título, optou-se por uma discretização da seguinte forma :

- 1 (menos um), se a relação for 4 ou mais vezes menor;
- 0,5 (menos meio), se a relação for 2 ou mais vezes menor;
- 0 (zero), se a relação for entre 2 vezes menor e 2 vezes maior;
- 0,5 (meio), se a relação for 2 ou mais vezes maior;
- 1 (um), se a relação for 4 ou mais vezes maior.

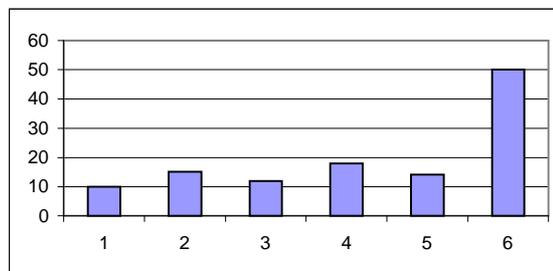


figura 3 - Exemplo de volume negociado

No gráfico acima (figura 3), os cinco dias anteriores apresentam uma média de 13,8 (1 a 5) e o dia atual (6) apresenta um valor de 50, ou seja, 3,6 vezes maior assim, a entrada será de 0.5.

1.1.3 Índice IBOVESPA

Corresponde ao IBOVESPA porém, foi saturado em 1 e -1. Ou seja, os valores não ultrapassam estes limites.

1.1.4 Arquitetura da Rede Neural

Ver figura 1 - Arquitetura da Rede.

4 O algoritmo de Aprendizado por Reforço – RN-AR

Foi utilizado o algoritmo linear gradiente descendente SARSA (λ) com política ϵ -Greedy (Ribeiro, 1999), (Sutton e Barto, 1998-pg.179 e 212)

$$\theta_{t+1} = \theta_t + \alpha [(r - V(s)) + \gamma V(s')] \gamma \lambda \in \quad (1)$$

Este algoritmo trabalha com entradas binárias referentes a estados, como só há um estado possível a cada iteração, somente uma sinapse é atualizada por vez. Para adequar o algoritmo às necessidades de atualização de todas as sinapses a cada iteração, este sofreu a seguinte modificação: a entrada de estados “ ϵ ” foi alterada de forma que, a atualização das sinapses seja feita com o “ ϵ ” correspondendo ao valor de entrada de cada sinapse no momento atual.

Como existem três neurônios, cada um responsável pela definição da utilidade de uma das ações, para acelerar o treinamento da rede neural realiza-se o treinamento no neurônio da ação escolhida e, caso esta ação não seja a indicada pelo algoritmo de escolha da ação ótima, treina-se também o neurônio da ação ótima.

4.1 Reforço

Para cada ação ótima escolhida aplica-se um reforço de +1 (um positivo) e para cada ação que não seja ótima aplica-se um reforço de -10 (dez negativo). A ação ótima é obtida através de um exame da movimentação futura.

4.2 Ação Ótima

Devido à complexidade de uma análise dos gráficos passo a passo, a ação ótima desejada (e por consequência o reforço) é calculada usando-se a média dos próximos 3 (três) dias e comparando-se esta como a do dia atual. Se o valor do dia atual for maior que a média, escolhe-se a ação de compra; se menor, escolhe-se a ação de venda e caso fique próximo da média com uma margem percentual pode-se assumir, manutenção.

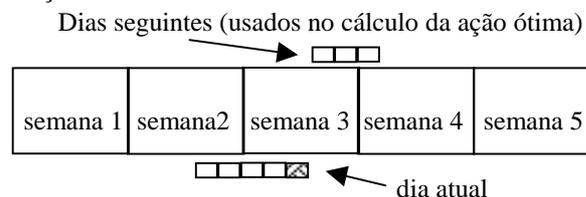


figura 4 – Escolha da ação ótima

Sendo assim, somente após o decorrer dos próximos 3 dias de pregão é que a rede será atualizada para o dia atual.

4.3 Treinamento da rede

Para o treinamento da rede, ou seja, o ato de atualizar os pesos das sinapse dos neurônios buscando que a ação escolhida pela rede seja a ação desejada, foram definidas as seguintes constantes:

$\alpha = 0,4$ (taxa de aprendizado)

$\gamma = 0,9$ (fator de desconto temporal)

$\lambda = 0,8$ (fator de desconto das diferenças temporais)

Demais variáveis usadas:

θ = valor da sinapse

ϵ = valor a entrar na sinapse

r = reforço

$V(s)$ = utilidade do estado atual

$V(s')$ = utilidade do estado futuro

O algoritmo usado para atualizar os neurônios:

Calcula-se a utilidade (valor de saída) de todos os neurônios (ações).

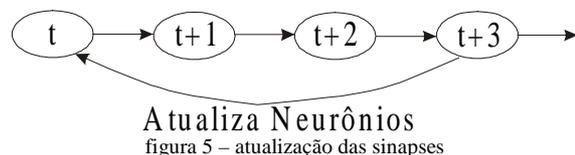
$$V(s) = \sum_{i=1}^I \theta_i * Entrada_i$$

A ação a ser tomada é a de maior utilidade.

Calcula-se a ação ótima (conforme item 4.2) e por consequência o reforço com relação a ação tomada.

Calcula-se a utilidade futura do neurônio a ser atualizado.

Aplica-se a Equação 1, para cada sinapse θ_i do neurônio, obtendo-se a nova sinapse.



5 Resultados

Foram realizadas experiências com os 54 ativos mais negociados do BOVESPA. Optou-se por estes ativos devido sua grande Presença (Número de pregões em que esteve presente/Número de pregões no período), o que permitiu uma análise mais abrangente, bem como um intervalo de dias de atuação sem grandes discontinuidades entre os ativos. Durante a simulação, considerou-se apenas o lucro gerado pela compra e posterior venda de ativos; não sendo agregado ganhos do tipo: dividendos, bonificações ou subscrição. Também, só foram compradas quantias unitárias de cada título e não foi montado um portfólio dinâmico de ativos, o que permitiria um aumento no rendimento, pois enquanto alguns ativos estão em alta, outros podem estar em baixa, fazendo com que o portfólio tenha que ser redimensionado diariamente para aumentar o lucro.

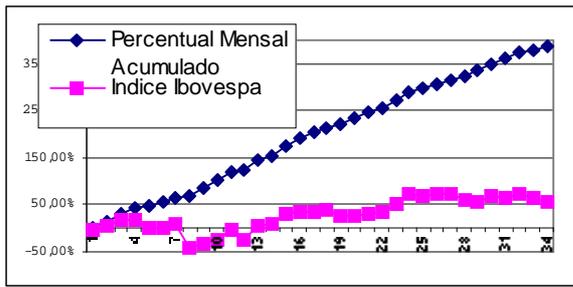


figura 6 - Rede Neural versus Ibovespa

Para o gráfico acima (figura 6), calculou-se o lucro mensal acumulado da carteira de ativos, no período de 01/1998 a 27/10./2000. O lucro acumulado da carteira atingiu o patamar de 388,31% enquanto o índice IBOVESPA atingiu 54,90%.

Nos gráficos seguintes (figura 7 e figura 8) (Haugen, 2000 - pg. 81), são demonstrados os resultados de grandes companhias mundiais de investimento em mercado de títulos, com resultados considerados excelentes; pode-se observar que no intervalo de aproximadamente 3 (três) anos, estas obtiveram lucros da margem de 119,89% e 99,12%; 66,49% e 49,33%. Analisando os dois gráficos, observa-se que um lucro total de 120% resulta, em média, um lucro de aproximadamente 30% ao ano; o que confirma o retorno prometido por operadoras nacionais para fundos de alto risco que “garantem” lucros de até 30% ao ano.

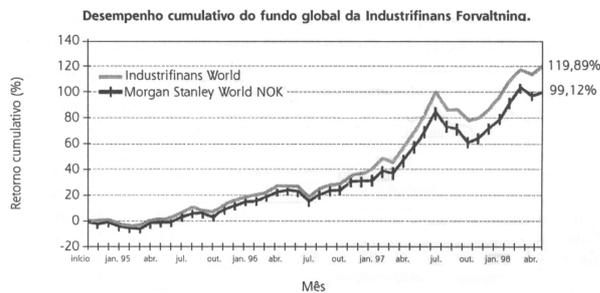


figura 7 - Fundo Industrifinans forvaltning (Haugen, 2000)

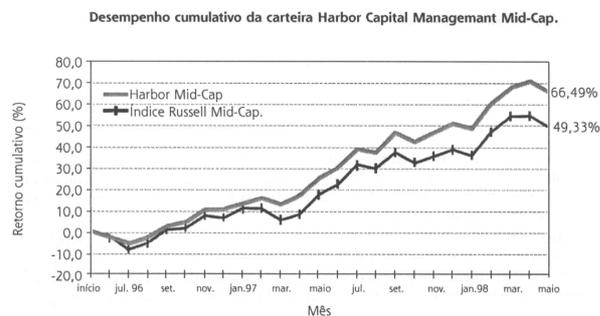


figura 8 - Carteira Harbor Capital M. Mid-Cap (Haugen, 2000)

Observando a tabela a seguir (Tabela 1), na qual estão os retornos de alguns dos ativos analisados pela RN, (foram selecionados ativos com representatividade para montar a tabela ilustrativa pois se fossem usados todos a tabela seria muito extensa), pode-se analisar o lucro no melhor caso, no pior e na média. Analisando o maior e o menor retorno em cada ano,

pode-se observar que no pior caso (o menor lucro foi de 27,7%), este retorno anual esteve próximo da faixa prometida de 30%, e na média, o retorno anual ficou acima de 120%, ou seja, um retorno quatro vezes superior ao prometido 30%.

Tabela 1 - Comparação entre diferentes ativos

	Ano1	Ano2	Ano3	Total
Acess4	225,42%	200,24%	91,61%	517,27%
CESP4	148,94%	142,09%	208,61%	499,65%
CNFB4	79,44%	164,92%	91,54%	335,90%
CPSL3	61,04%	73,85%	31,46%	166,36%
ELPL4	88,20%	261,89%	123,34%	473,44%
LAME4	210,56%	214,87%	120,08%	545,51%
MAGS5	27,69%	72,84%	66,13%	166,67%
PCAR4	82,38%	141,42%	55,95%	279,75%
PETR3	82,99%	183,01%	128,92%	394,93%
TGRM6	111,54%	79,24%	86,79%	277,57%
USIM5	143,94%	274,88%	149,30%	568,13%

Considerando todos os ativos				
Média de todos os ativos				
	125,31%	166,19%	97,60%	391,47%
Max	269,81%	274,88%	208,61%	568,13%
Min	27,69%	53,67%	31,46%	166,36%

Obs. O terceiro ano vai de 01/01/2000 a 27/10/200; ou seja, ainda faltam dois meses para completar o ano.

6 Comparação com rede Perceptron

Usando a mesma configuração da rede anterior, trocou-se apenas o algoritmo de Aprendizado por Reforço por um algoritmo tipo *Perceptron* (Nascimento e Yoneyama, 2000).

Como pode ser visto na tabela a seguir, a rede perceptron para o melhor título (apenas o título de maior rendimento), obteve um retorno inferior ao da RN-AR e para melhorar a predição a Rede Perceptron precisou ser treinada off-line, diferente da RN-AR que teve diferenças mínimas no caso de treinamento off-line. Para a RN-AR, o retorno esperado foi de 517,27% enquanto para a Perceptron o maior retorno foi de 220,11%. Enfatizando que, para outros ativos, o treinamento necessário para a rede Perceptron obter seu valor máximo, diferiu do 1º treinamento. Para os outros ativos, a diferença do retorno das duas arquiteturas variou em média 93% a favor da RN-AR.

Além da rede Perceptron não conseguir, em muitos casos, obter um lucro cumulativo acima de 80% para os últimos três anos, a quantidade de treinamentos para que a rede obtenha o maior lucro diferiu substancialmente de título para título (Tabela 2).

Tabela 2 - Resultados da Rede Perceptron

Tipo de treinamento usado na rede Perceptron	Índice acumulado dos últimos 3 anos
durante 1º treinamento	82,44298135
já treinada, com 1º treinamento seguinte	220,110729
durante 6º treinamento	187,7547466
Já treinada, com 6º treinamento seguinte	175,6724893
durante 20º treinamento	41,55506386
já treinada, com 20º treinamento seguinte	73,23473288

7 Outras redes previsoras

Apesar dos vários estudos em andamento e aqueles que já se tornaram produtos comerciais para previsão do mercado acionário, vale ressaltar algumas diferenças marcantes entre as redes existentes e a proposta: existe uma grande diferença na quantidade de entradas, as redes usuais normalmente possuem um conjunto de 270 entradas referentes ao movimento dos últimos 30 (trinta) dias e ainda alguns índices econômicos. Na rede proposta, foi usado apenas um conjunto restrito de 17 (dezessete) dados. A rede proposta possui apenas 3 (três) neurônios, enquanto as demais, utilizam uma arquitetura mais complexa (camadas escondidas).

8 Conclusão

Observa-se que, não foram realizadas comparações com redes neurais com camadas escondidas, pois o objetivo deste trabalho, foi provar que uma rede neural com Aprendizado por Reforço (AR), poderia ser usada para previsões temporais uso ainda desconhecido pelos autores deste paper.

A implementação do algoritmo utilizando a Rede Neural com Aprendizado por Reforço para a previsão da movimentação do mercado de ativos e possível negociação foi bem sucedida, apesar da sua simplicidade, usando apenas 17 (dezessete) dados de entrada e uma arquitetura de rede extremamente simples, com apenas 3 neurônios de entrada, nenhuma camada escondida e os neurônios de saída sendo os de entrada.

As negociações experimentais feitas com o algoritmo Perceptron trouxeram um lucro menor e, em algumas vezes, prejuízo. Isto se deve, provavelmente, à ausência de camadas escondidas, que, se presentes, levariam a uma melhora do poder de aproximação, representando melhor o modelo.

Comparando o retorno da Rede Neural proposta com investimentos reais de empresas que realizam operações no mercado mundial, e que obtêm grandes lucros, estes acima de 30% ao ano, observa-se a grande diferença nos lucros obtidos pelo agente pro-

posto, pois ele resultou, em média, um retorno anual acima de 110%.

Como próximo passo, a implementação de uma RNAR com camadas escondidas possibilitaria a comparação desta arquitetura com as atualmente usadas no mercado e portanto a seleção da melhor rede.

Agradecimentos

Os autores gostariam de agradecer o apoio da Fundação Aplicações de Tecnologias Críticas – ATECH que em conjunto com o SIVAM possibilitaram a realização deste projeto.

Referências

- Blum, Adam (1992). Neural Networks in C++; An Object-Oriented Framework for Building Connectionist Systems; John Wiley & Sons, Inc.
- Haugen, A. Robert (2000). Os Segredos da Bolsa, Prentice Hall.
- Nascimento, Cairo L. Jr. e Yoneyama, Takashi (2000). Inteligência Artificial em Controle e Automação, Ed. Edgard Blücher LTDA.
- Ribeiro, Carlos H. C. (1999). Aprendizado por reforço, V Escola de Redes Neurais, CBRN1999.
- Sutton, S. Richard; Barto, Andrew G. (1998). Reinforcement Learning - An Introduction, Mit Press.

Internet

- Site do investshop “www.investshop.com.br”
- Site com a Base de Dados da Bolsa de Valores “www.gripper.hpg.com.br/download.html”