

Previsão de Tendência de Preços de Ações: Uma Tentativa com a WiSARD

Cleiton Moya de Almeida¹

¹Programa de Engenharia de Computação e Sistemas (PESC)
Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia (COPPE)
Universidade Federal do Rio de Janeiro (UFRJ)

cma@cos.ufrj.br

Resumo. Neste trabalho, o modelo de rede neural sem peso WiSARD (Wilkie, Stonham & Aleksander's Recognition Device) foi testado na tarefa de previsão da tendência de preços de ações negociadas na bolsa de valores brasileira. Como este modelo utiliza dados de entrada no formato binário, foi proposta a construção de uma retina utilizando sinais binários gerados por indicadores da Análise Técnica. O problema abordado foi o de classificação da tendência (direção do preço) para um horizonte de “h” dias seguintes ao dia atual. Os experimentos apresentaram acurácia não satisfatória para a classificação de tendências. O modelo foi comparado com um semelhante na literatura, e então algumas hipóteses foram levantadas para a baixa acurácia obtida.

1. Introdução

1.1. Fundos quantitativos e aprendizado de máquina

Os fundos de investimentos quantitativos utilizam técnicas estatísticas e computacionais para a tomada de decisão e negociações automatizadas. Em 2017, segundo o Morgan Stanley, os fundos quantitativos movimentaram US\$ 1,5 trilhões e, em relação aos 6 últimos anos, tiveram crescimento anual de 15%. O *Medallion Fund*, um dos pioneiros na área, acumula de 1988 a 2018 um retorno anual médio de 39% (Jansen, 2018).

Nos últimos anos, tem crescido o interesse de aplicação de técnicas de Aprendizado de Máquina na área financeira. Morgan Stanley relata que apenas 23% de seus clientes de fundos quantitativos não utilizam ou não consideram utilizar Aprendizado de Máquina em seus fundos (Jansen, 2018).

A aplicação de técnicas de aprendizado de máquina na área financeira apresenta bastante desafios devido à natureza complexa da tarefa. Abu-Mostafa & Atiya (1996) exemplificam que, enquanto na tarefa de reconhecimento de caracteres as predições concordam $1 - \epsilon$ das vezes, na área financeira as predições concordam $\frac{1}{2} + \epsilon$ vezes (performance na faixa de 50%). Atualmente, embora o estado arte reporte resultados com acurácia superior a 80% (Henrique, Sobreiro e Kimura, 2019), Prado (2018) comenta que, na prática, a maioria dos algoritmos fracassam. Desta forma, este é um campo com bastante oportunidades para o desenvolvimento de trabalhos.

1.2. WiSARD

Dentre os diversos modelos para aprendizagem de máquina, a rede neural sem peso WiSARD (*Wilkie, Stonham & Aleksander's Recognition Device*) (Alexander et al., 1984) mostra-se bastante atrativa para a área de negociação algorítmica, pois permite o treinamento de forma *on-line* e com uma única iteração, ao contrário do treinamento em batelada e múltiplas iterações requeridas por outros modelos.

Utilizando a WiSARD, Alves (2017) mostrou ser viável a predição de tendência de ações através de uma classificação de tendências em 6 fases (Recuperação, Acumulação, Altista, Aviso, Distribuição, Baixista), calculadas com base no preço de fechamento. Entretanto, neste modelo, a retina da rede considera apenas o preço de fechamento da ação; este preço é usado para o cálculo de duas médias móveis, que por sua vez determinam qual é a fase de tendência. Além disso, conforme mostrado no trabalho de Alves (2017), esta classificação não se mostra adequada para utilização em estratégias de negociações simples, requerendo políticas mais elaboradas.

Por outro lado, é comum na área de Análise Técnica e em trabalhos de Aprendizado de Máquina aplicados ao mercado financeiro a utilização de outros dados de mercado (preço de abertura, preço máximo, mínimo e volume de negociações). É comum a utilizar ainda, como atributos (*features*), indicadores derivados, tais como indicadores da Análise Fundamentalista ou Análise Técnica. Nota-se ainda que a maioria os trabalhos abordam o problema de classificação de tendências utilizando apenas duas classes: (tendência de alta e tendência de baixa, rotuladas à partir do preço ou retorno (Atsalakis & Valavanis, 2009).

1. Objetivo do trabalho

O objetivo deste trabalho é avaliar a aplicação da rede WiSARD na tarefa a classificação de tendência de ações utilizando uma retina construída com sinais binários gerados a partir de indicadores da Análise Técnica, tais como médias móveis, Bandas de Bollinger e Índice de Forças Relativas (*Relative Strenght Index* - RSI).

1.1. Base de Dados

Os dados de mercado (preços de abertura, máxima do dia, mínima do dia, fechamento e volume de negociação) foram extraídos da plataforma comercial *Profit Clear Trade*®, desenvolvida pela empresa Nelógica e distribuído pela corretora Clear. (Clear, 2020). Ou seja, utilizou-se uma série temporal (período diário), multidimensional, para cada ativo testado. Conforme explicado nas seções seguintes, os seguintes períodos foram utilizados: (i) 01/01/2008 a 31/06/2015 e (ii) 01/0/2015 a 31/01/2020.

Foram utilizados ainda dados de 3 ativos: BBDC4 (ações preferenciais do Bradesco), PETR4 (ações preferenciais da Petrobras) e VALE3 (ações da Vale), a fim de comparação de resultados com o trabalho de Alves (2017).

2. Modelagem

2.1. Atributos - indicadores técnicos e sinais binários

Para a criação da retina, foram utilizados de 9 indicadores da análise técnica, mostrados na Tabela 1 Tabela 2. Uma apresentação detalhada destes indicadores é feita em (Lemos,

2018). Estes 11 indicadores geraram 21 parâmetros de entrada, também mostrados na Tabela 1.

Com base nestes indicadores e nos dados de mercado, foram construídos 28 sinais binários (Tabela 2).

Os indicadores forem implementados de forma que eles podem ser selecionados e usados de forma individual em cada experimento, permitindo múltiplas combinações e cenários.

Tabela 1: Indicadores técnicos e seus parâmetros de entrada

Indicador	Classificação (Lemos, 2018)	Parâmetro	Valor inicial
Médias móveis	Seguidor de tendência	Período da média móvel curta (ma_s)	10
		Período da média móvel longa (ma_L)	50
Bandas de Bollinger	Seguido de tendência	Período (T_{BOL})	20
		Número de desvios-padrões (BOL_σ)	2
ADX (<i>Average Directional Moviment Index</i>)	Seguidor de tendência	Período (T_{ADX})	14
		Limite superior de força (ADX_{UL})	25
Aroon	Seguidor de tendência	Período (T_{AROON})	14
		Limite de força ($AROON_{SL}$)	80
		Limite de fraqueza ($AROON_{WL}$)	40
Canal Donchian	Seguidor de tendência	Período da banda superior (T_{SUP}^{DON})	20
		Período da banda inferior (T_{INF}^{DON})	20
Canal de Keltner	Seguidor de tendência	Período (T_{KELT})	20
MACD (<i>Moving Average Convergence-Divergence</i>)	Oscilador de <i>momentum</i>	Período da média móvel curta (T_{SMA}^{MACD})	12
		Período da média móvel longa (T_{LMA}^{MACD})	26
		Período do “sinal” (T_{SIG}^{MACD})	9
RSI (<i>Relative Strength Index</i>)	Oscilador de <i>momentum</i>	Período (T_{RSI})	14
		Limite inferior de sobrecompra (RSI_{LL})	30
		Limite superior sobre venda (RSI_{UL})	70
MFI (<i>Money Flow Index</i>)	Oscilador de volume	Período (T_{MFI})	14
		Limite superior (MFI_{SUP})	80
		Limite inferior (MFI_{INF})	20

Tabela 2: Sinais binários utilizados para a construção da retina

Denominação do sinal binário	Dado de mercado ou indicador técnico utilizado	Critério para a geração do sinal
Volume crescente	Volume financeiro	Volume crescente nos 3 últimos períodos: $vol(t) > 1,1 \times vol(t-1)$ $> 1,1 \times vol(t-2)$
Volume decrescente	Volume financeiro	Volume decrescente nos 3 últimos períodos: $vol < 0,9 \times vol(t-1)$ $< 0,9 \times vol(t-2)$
Preços crescente	Preço de fechamento	Preço crescente nos 3 últimos períodos: $P_t > 1,005 P_{t-1} > 1,005 P_{t-2}$
Preços decrescente	Preço de fechamento	Preço decrescente nos 3 últimos períodos: $P_t < 0,995 P_{t-1} < 0,995 P_{t-2}$
Médias móveis – tendência de alta	Médias móveis	Média móvel curta acima da média móvel longa: $ma_s > ma_L$
Cruzamento de médias – compra	Médias móveis	Média curta cruza a média longa de baixo para cima: $ma_s(t) > ma_L(t)$ AND $ma_s(t-1) < ma_L(t-1)$
Cruzamento de médias – venda	Médias móveis	Média curta cruza a média longa de cima para baixo: $ma_s(t) < ma_L(t)$ AND $ma_s(t-1) > ma_L(t-1)$
Banda de Bollinger – sobrecompra	Preço de fechamento / Banda de Bollinger	Preço acima da banda superior
Banda de Bollinger – sobre venda	Preço de fechamento / Banda de Bollinger	Preço abaixo da banda inferior
MACD – compra	MACD	Cruzamento de baixo para cima da linha ‘macd’ com a linha “signal” do indicador.

Denominação do sinal binário	Dado de mercado ou indicador técnico utilizado	Critério para a geração do sinal
MACD – venda	MACD	Cruzamento de cima para baixo da linha “macd” com a linha “sinal”
MACD – tendência de alta	MACD	Linha “sinal” acima da linha “macd”
RSI – sobrecompra	RSI	$RSI > RSI_{UL}$
RSI – sobrevenda	RSI	$RSI < RSI_{LL}$
Donchian Channel – sobrecompra	Preço de fechamento / Canal Donchian	Preço acima da banda superior
Donchian Channel – sobrevenda	Preço de fechamento / Canal Donchian	Preço abaixo da banda inferior
Keltner Channel – sobrecompra	Preço de fechamento / Canal de Keltner	Preço acima da banda superior
Keltner Channel – sobrevenda	Preço de fechamento / Canal de Keltner	Preço abaixo da banda inferior
ADX – tendência com força	ADX	$ADX > ADX_{UL}$
ADX – compra	ADX	$DI+$ cruza $DI-$ de baixo para cima.
ADX – venda	ADX	$DI+$ cruza $DI-$ de cima para baixo.
ADX – tendência de alta	ADX	$DI+$ está acima de $DI-$
Aroon – força na curva “para cima”	Aroon	Curva “para cima” acima de $AROON_{SL}$
Aroon – força na curva “para baixo”	Aroon	Curva “para baixo” acima de $AROON_{SL}$
Aroon – fraqueza na curva “para cima”	Aroon	Curva “para cima” abaixo de $AROON_{WL}$
Aroon – fraqueza na curva “para baixo”	Aroon	Curva “para baixo” abaixo de $AROON_{WL}$
MFI – compra	MFI	$MFI > MFI_{UL}$
MFI – venda	MFI	$MFI < MFI_{LL}$

2.2. Janela deslizante e construção da retina

A partir dos sinais binários, a série temporal foi representada por um vetor de estados com o valor de cada sinal s_m no instante de tempo t :

$$x_t = [s_1, s_2, \dots, s_M].$$

A retina então foi construída utilizando uma sequência com T períodos, incluindo o período atual:

$$X_t = [x_t, x_{t-1}, \dots, x_{t-T+1}].$$

Desta forma, o tamanho da retina é dado por $N = N_s \times T$, onde N_s é o número de sinais utilizados (tamanho do vetor x_t).

2.3. Rotulagem dos dados

Para a rotulagem dos dados, utilizou-se a seguinte regra, considerando o preço no período atual (P_t) e o preço no horizonte h (P_{t+h}):

- $P_{t+h} < P_t$: tendência de baixa (rótulo “0”);
- $P_{t+h} \geq P_t$: tendência de alta (rótulo “1”).

3. Experimentos e resultados

3.1. Recursos utilizados

Os experimentos foram simulados em um computador pessoal com processador Intel Core™ i7-4510U (4x2.00Ghz), 7,7 GiB de memória RAM e sistema operacional Ubuntu 20.04 LTS.

Os algoritmos foram implementados na linguagem Python 3.7. Utilizou-se se a biblioteca wisardpkg v1.6.3 (Filho et al, 2020) para implementação do modelo WiSARD.

As séries dos indicadores técnicos foram calculadas com a biblioteca Pandas TA (Jhonson, 2020).

3.2. Parâmetros avaliados

As execuções foram executadas inicialmente com os parâmetros listados na Tabela 1. As seguintes combinações de parâmetros adicionais foram avaliadas:

- Número de bits de endereçamento das RAMs: [4, 8, 16, 24];
- Número de períodos da janela móvel (T): [1, 3, 5, 15];
- Horizonte (h): [1, 3, 5, 10].

3.3. Períodos e divisões do conjunto de dados

A fim de comparar os resultados deste trabalho com (Alves, 2017), bem como avaliar o desempenho em um período mais recente, os seguintes períodos foram utilizados para as séries temporais:

- 01/01/2018 a 31/06/2015;
- 01/01/2015 a 31/01/2020.

As séries foram então divididas sequencialmente em 3 conjuntos de dados: 1 para treinamento (60%), uma para testes (20%) e outra para validação (20%).

3.4. Resultados – Previsão de tendências

Em todos os cenários testados, observou-se uma elevada variabilidade na acurácia média dos experimentos. Esta variabilidade deve-se ao mapeamento pseudo-aleatório dos *bits* da entrada nas RAMs da WiSARD, no momento de criação do modelo (Filho, 2020).

A fim de melhor quantificar esta variabilidade, como os tempos médios de cada experimento foram da ordem de 20 milissegundos, executou-se então um teste com 1000 experimentos para cada cenário, calculando-se a acurácia média e o desvio padrão. Porém, em todos estes testes, a acurácia média obtida foi em torno de 50%, dentro do limite de erro do desvio padrão, conforme mostrado na Tabela 3. Sendo assim, os resultados de acurácia não foram satisfatórios.

Tabela 3: Acurácia com a WiSARD para a previsão de tendências

	Alves, 2017	Este trabalho 2008 a 2015	Este trabalho 2015 a 2020
BBDC4	40,21%	(50,9 ± 2,2) %	(53,6 ± 2,6) %
PETR4	65,26%	(51,8 ± 2,0) %	(49,5 ± 3,5) %
VALE3	55,09%	(52,0 ± 2,3) %	(51,2 ± 2,7) %

3.5. Análise e comparação de resultados

A Tabela 3 compara os resultados obtidos neste trabalho com alguns resultados obtidos por (Alves, 2017). Nesta tabela, reportou-se apenas os resultados dos ativos BBDC4, PETR4 e VALE4; porém, convém observar que Alves (2017) também realizou experimentos com outros ativos. Pela tabela, é possível concluir que a acurácia obtida neste trabalho foi inferior aos resultados obtidos por (Alves, 2017) para os ativos PETR4 e VALE3. O ativo BBDC4 apresentou maior acurácia no presente trabalho, porém esta acurácia não é satisfatória devido ao seu desvio padrão, conforme comentado na seção anterior.

Para analisar os motivos desta diferença de desempenho, primeiramente é importante notar que Alves (2017) utilizou seis classes diferentes para classificação,

enquanto este trabalho utilizou apenas duas. Em princípio, a utilização de apenas duas classes pode ter contribuído, neste trabalho, para uma maior ocorrência de empates na classificação, degradando desta forma a acurácia.

Porém, talvez o principal motivo para a diferença esteja no método de rotulagem dos dados. Em (Alves, 2017), os rótulos determinados com base nas fases calculadas pelas médias móveis. Ou seja, a variável predita, naquele trabalho, pode ser considerada um indicador técnico, e não o estado real do ativo. Por exemplo, nesta metodologia, é possível a fase atual ser determinada como “altista”, a fase seguinte (predita) também ser classificada “altista” e mesmo assim o preço na fase seguinte ser menor que o preço atual. Isto pode ser observado na Figura 1.

Já este trabalho, a rotulagem foi realizada não com base em indicadores, mas sim com base na diferença real de preços entre a data atual e a futura. Assim, a saída “tendência de alta” sempre está associada um real um aumento da cotação. É fácil ver então que a variável predita no presente trabalho é mais complexa que a variável utilizada por Alves (2017), pois o real estado da tendência possui maior variação em um intervalo de tempo Δt do que a variação das fases calculadas com base nas médias móveis, como ilustrado na Figura 2.

Devido à complexidade do sinal de predição (Figura 2), uma hipótese para a baixa acurácia verificada neste trabalho é que os sinais binários utilizados não contem o grau de informação necessária para a predição requerida. Neste sentido, um possível estudo futuro seria verificar o quanto de informação cada sinal individual agrega ao modelo, bem como avaliar a inclusão de mais informação na retina. Outra possibilidade seria, ao invés de usar sinais binários, discretizar o valor dos indicadores (número real).

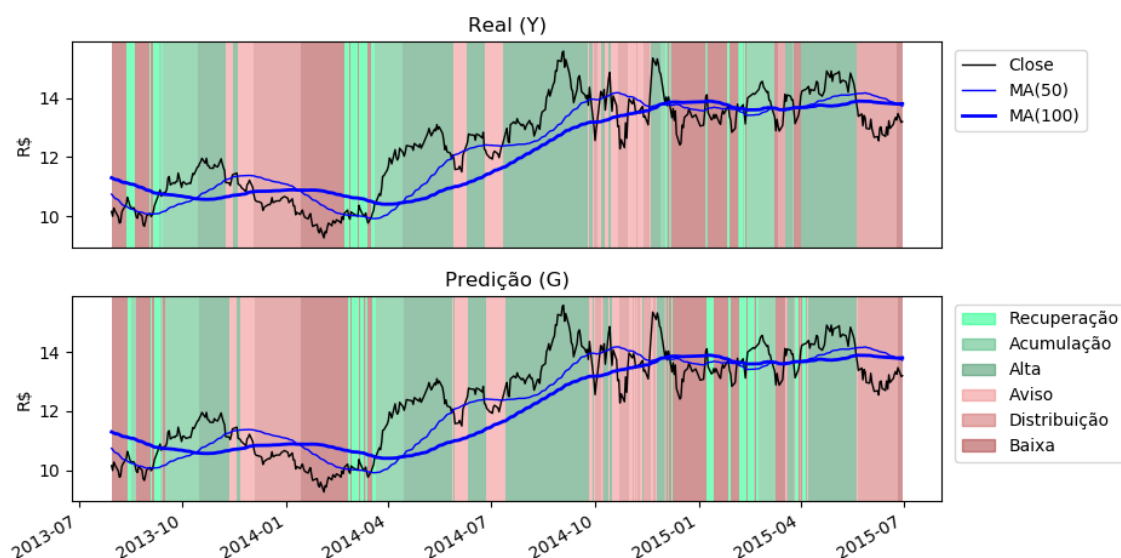


Figura 1: Fases calculadas x previstas - metodologia de (Alves, 2017)
Fonte: elaborado pelo autor

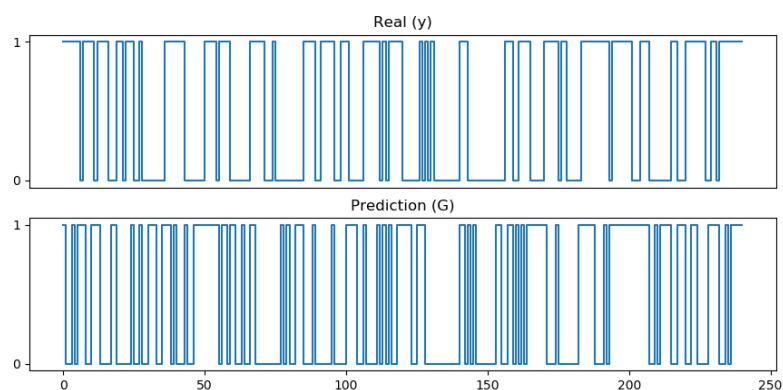


Figura 2: Tendência real x predita – metodologia deste trabalho

4. Conclusões

Neste trabalho a rede WiSARD foi testada para a predição da tendência de preços de ações no mercado brasileiro utilizando como *features* sinais binários gerados a partir de indicadores da análise técnica. Foram simuladas diversas configurações de parâmetros. Entretanto, nenhuma delas foi capaz de apresentar predições com uma acurácia aceitável.

Um provável motivo para a baixa acurácia é que os sinais utilizados, por serem binários, não contem o grau de informação requerido para a predição. Neste caso, uma sugestão de trabalho futuro seria fazer uma análise individual de cada *feature*, do quanto ela é capaz de “agregar” informação ao modelo. Outra possibilidade seria avaliar a discretização dos valores dos indicadores (número real).

Referências

- Abu-Mostafa, Y. S., Atiya, A. F. (1996). Introduction to financial forecasting. *Applied Intelligence*, 6 (3), 205–213
- Alves, Samara. A. *Negociação no mercado financeiro utilizando a rede neural sem peso WiSARD*. Dissertação de mestrado. 2017.
- Atsalakis, G. S., Valavanis, K. P. (2009). Surveying stock market forecasting techniques–Part II: Soft computing methods. *Expert Systems with Applications*, 36 (3), 5932–5941.
- Clear. *Profit trader Clear*. Disponível em: <<https://www.clear.com.br/site/plataformas/profit-trader-clear>>. Acesso em: 16/07/2020.
- Filho, A. S. L., Guarisa, G. P., Leopoldo Lusquino Filho, Luiz F. R. Oliveira, Felipe M. G. França, Priscila M. V. Lima. *wisardpkg - A library for WiSARD-based models*“. CoRR abs/2005.00887 (2020).
- Henrique, B., Sobreiro, V., and Kimura, H. Literature Review: Machine Learning Techniques Applied to Financial Market Prediction. In: *Expert Systems with Applications* 124 (June 2019). DOI:10.1016/j.eswa.2019.01.012
- I. Aleksander, W. Thomas, and P. Bowden, Wisard a radical step forward in image recognition, *Sensor review*, vol. 4, no. 3, pp. 120–124, 1984. Knuth, D. E. (1984), The TeXbook, Addison Wesley, 15th edition.
- Jansen, S. *Hands-On Machine Learning for Algorithmic Trading*. Pocket: 2018
- Johnson, K. *Technical Analysis Library in Python 3.7*. Disponível em: <<https://twopirllc.github.io/pandas-ta/>>. Acesso em: 16/07/2020.
- Lemos, F. *Análise Técnica dos Mercado Financeiros*. Editora Saraiva Educação: 2018.
- Prado, M. L. *Advances in Financial Machine Learning*. Wiley, 2018.