

APRENDIZADO DE MÁQUINA APLICADO À DETECÇÃO DE TENDÊNCIAS EM SÉRIES TEMPORAIS FINANCEIRAS

MARCELO RUBENS GIESEL¹; RICARDO MATSUMURA ARAUJO²

¹Universidade Federal de Pelotas (UFPEL) – mgiesel@inf.ufpel.edu.br

²Universidade Federal de Pelotas (UFPEL) – ricardo@inf.ufpel.edu.br

1. INTRODUÇÃO

Uma empresa, ao abrir seu capital em uma bolsa de valores, divide sua sociedade em várias frações de igual valor, e cada fração é chamada de ação. O preço da ação de uma empresa sofre alterações devido ao processo de compra e venda, estimulado por decisões baseadas em fatores diversos. Estas alterações de preço oferecem oportunidades de negócios. Com a finalidade de obtenção de lucro, os operadores deste mercado fazem uso de estratégias variadas de aplicações para longo, médio ou curto prazo. Uma destas estratégias é a análise técnica, que faz uso da variação de preço de um ativo (ação) ao longo do tempo. Seguindo um modelo, um padrão na série temporal dos preços, os operadores decidem se realizam operações de compra ou venda acreditando em um movimento. Por exemplo, se acreditam que haverá um aumento no preço dos ativos eles compram para depois poder vender quando o preço estiver mais elevado, lucrando com a diferença.

Uma movimentação de preço com uma certa duração é uma tendência, e alguns modelos tentam identificar a ocorrência destes fenômenos. Existem identificadores de tendências, como médias móveis, canais de Donchian e HiLo Activator (MALHEIROS, 2011), que foram criados para visualizar melhor a formação de tendências e detectá-las utilizando diferentes critérios. Entretanto, ainda são verificados muitos falsos positivos, situações onde o indicador técnico sinaliza a existência de uma tendência mas que acaba não se confirmando. JOHANN (2010), por exemplo, obteve um percentual de operações lucrativas de 46% com a sua estratégia mais rentável de médias móveis. Essa baixa taxa de acerto faz com que os lucros obtidos sejam consumidos pelas operações negativas, o que motiva a busca por ferramentas mais precisas.

Recentemente, um novo algoritmo para detecção de tendências foi proposto. Sabendo da importância da detecção, classificação e predição de eventos na extração de informação útil de dados temporais, NIKOLOV (2012) estudou o problema de predição em um sistema complexo utilizando uma grande quantidade de dados. Focou especificamente na classificação binária de séries temporais, onde procurava separar “eventos” de “não-eventos” dada uma quantidade suficiente de exemplos históricos. Para isso elaborou um algoritmo de aprendizagem de máquina sem modelo para fazer inferências na série temporal. Na classificação, após transformar a série temporal em sinais, o sinal observado era comparado com dois conjuntos de sinais de referências, um contendo exemplos positivos de eventos e outro contendo exemplos negativos, de “não-eventos”. Assumindo que todos os sinais observados são uma forma ruidosa de um sinal ou do conjunto positivo ou do conjunto negativo, o sinal observado era comparado com os sinais dos dois conjuntos e era calculada a probabilidade de este sinal pertencer a um conjunto ou a outro. A comparação era feita calculando-se a distância do sinal observado para os sinais das classes de referência, permitindo assim que o sinal fosse classificado de uma forma não-paramétrica

diretamente a partir dos dados em si, sem a necessidade de especificar uma estrutura de modelo.

Para testar sua eficiência, a técnica foi aplicada na identificação de tendências para tópicos do Twitter. Em 95% das vezes conseguiu acertar quais tópicos seriam classificados como *trend topics*, a lista dos tópicos mais comentados durante um intervalo de tempo. Além de acertar na grande maioria das vezes, a identificação era realizada com uma média de aproximadamente uma hora e meia antes do tópico ser identificado como tendência pelos algoritmos da própria rede social. No mesmo trabalho NIKOLOV (2012) sugere que seu algoritmo poderia servir também para detecção de tendências no mercado financeiro, mas ele não realizou este teste.

A similaridade nas séries temporais, e a própria afirmação do autor da dissertação, nos levou a crer que o método de detecção de tendências apresentado pode ter aplicação no mercado de ações. Se isso se confirmar, mesmo uma taxa de acerto substancialmente inferior aos 95% verificados na detecção de tendências na rede social seria útil neste novo cenário, onde é comum encontrar taxas menores de acerto (JOHANN, 2010). Qualquer aumento na porcentagem de operações positivas representaria um aumento considerável no lucro, uma vez que teriam menos operações negativas e isso consumiria menos os recursos financeiros do operador.

Como o uso de técnicas de inteligência artificial têm se mostrado apreciável em séries temporais financeiras (GRILLI; SFRECOLA, 2009) (GRILLI et al., 2009), este trabalho tem como objetivo principal aplicar a técnica de aprendizado de máquina de NIKOLOV (2012) nas séries temporais do mercado de ações brasileiro e comparar os resultados com outra técnica tradicional amplamente utilizada neste mercado.

2. METODOLOGIA

A primeira parte do estudo consistiu na obtenção da base histórica dos preços de algumas das ações mais negociadas na bolsa de valores brasileira. Os dados foram obtidos a partir de um grande portal financeiro, o Yahoo Finance, que disponibiliza gratuitamente uma parte destes dados. Com os dados em mãos, foi selecionado um período histórico para análise da série temporal, e dividida em um conjunto de treinamento e outro de testes. O conjunto de treinamento foi utilizado para a criação dos conjuntos de exemplos positivos e negativos da técnica de NIKOLOV (2012). O conjunto de testes foi utilizado na execução da técnica, a fim de verificar sua capacidade de detecção das tendências.

O algoritmo de NIKOLOV (2012) foi então implementado com as mesmas configurações do trabalho original. Idealmente, pretendíamos fazer uma validação desta implementação, tentando reproduzir os resultados obtidos naquele trabalho. Entretanto, tal validação não foi possível de ser realizada, pois o autor não tem permissão para divulgar os dados utilizados em seus testes. Desta forma a validação foi realizada de outra maneira, utilizando os dados já obtidos das ações.

A fase atual é a de adaptação do algoritmo à detecção de tendências na bolsa de valores. O algoritmo possui uma grande quantidade de variáveis que influenciam no resultado, e devem ser ajustadas a cada teste para convergir para a obtenção de melhores resultados.

O passo seguinte será a otimização da técnica convencional para futura comparação com o algoritmo implementado. A técnica convencional será

otimizada baseada no mesmo conjunto de treinamento e executada sobre o mesmo conjunto de testes que o algoritmo.

3. RESULTADOS E DISCUSSÃO

A validação do algoritmo foi feita utilizando os próprios dados obtidos das ações. O conjunto de treinamento foi utilizado no treinamento do algoritmo e também como conjunto de testes. Nesta configuração o algoritmo deveria reconhecer todas as tendências positivas contidas no conjunto de treinamento para estar correto. Este foi o resultado obtido, garantindo o funcionamento do algoritmo e mostrando que é capaz de reconhecer os sinais apresentados por este ambiente.

Os primeiros resultados mostram que o algoritmo é capaz de identificar tendências ainda em seu início, proporcionando um bom aproveitamento deste movimento. Permitir que a operação seja iniciada com os preços ainda baixos significa uma possibilidade maior de lucro na operação. Entretanto esta detecção precoce também acarreta em um número indesejado de falsos positivos, quando o algoritmo indica erroneamente o início de uma tendência. Buscando a diminuição destes resultados negativos.

A fase atual é a de adaptação do algoritmo à detecção de tendências, e esta etapa requer alguns cuidados. Na rede social, a tendência está relacionada a comportamentos sociais e a série temporal diz respeito unicamente ao volume de mensagens com o tópico específico. Na bolsa de valores, a série temporal diz respeito não ao volume negociado do ativo, mas ao seu preço, e a tendência pode ser formada também por outros fatores, como balanços empresariais e políticas governamentais. Ainda, a identificação de tendências para formação dos conjuntos de exemplos negativos e positivos é feita de forma mais simples na rede social. O próprio Twitter oferece a informação dos tópicos que foram identificados como tendências, facilitando a criação do conjunto de exemplos positivos. Na bolsa de valores não temos tal auxílio. É preciso encontrar uma forma que defina com clareza o que é uma tendência e selecionar exemplos para a criação do conjunto dos exemplos positivos. Todos estes pontos devem ser tratados com especial atenção durante a adaptação do algoritmo.

4. CONCLUSÕES

Este artigo apresentou os resultados preliminares e algumas das discussões da aplicação de um método de aprendizado de máquina à detecção de tendências em séries temporais financeiras.

O algoritmo foi implementado e foi capaz de reconhecer tendências que lhe foram apresentadas previamente. Isso mostra que pode lidar com séries temporais com estas características. As fases seguintes, compostas pela otimização do algoritmo e pela comparação de seus resultados com outras técnicas, apresentarão mais conclusões a respeito das vantagens de seu uso.

5. REFERÊNCIAS BIBLIOGRÁFICAS

GRILLI, L.; RUSSO, M.A.; SFRECOLA, A. Financial time series and neural networks in a minority game context. **Mathematical and Statistical Methods for Actuarial Sciences and Finance**, Springer Milan, p. 153-162, 2010.

GRILLI, L.; SFRECOLA, A. A neural networks approach to minority game. **Neural Computing and Applications**, v.18, n.2, p. 109-113, 2009.

JOHANN, F.J.; VIDAL, J.P.; TÁRTARO, M. **Utilização dos indicadores médias móveis para a análise gráfica de ações**. 2010. 94f. Monografia (Bacharelado em Administração) - Curso de Graduação em Administração, Faculdade Assis Gurgacz.

MALHEIROS, R.S. **Operando com trading systems na bolsa de valores**. São Paulo: Évora, 2011.

NIKOLOV, S. **Trend or no trend: a novel nonparametric method for classifying time series**. 2012. 68f. Dissertação (Mestrado em Engenharia Elétrica e Ciência da Computação) - Curso de Pós-graduação em Engenharia Elétrica e Ciência da Computação, Massachusetts Institute of Technology.