

Uma Aplicação de Meta-Aprendizagem no Mercado Euro/Dólar

Felipe de Almeida Duarte

Renato Cesar Sato

Ana Carolina Lorena

Universidade Federal de São Paulo
Instituto Tecnológico Aeronáutico
Programa de Pós-Graduação em Pesquisa Operacional

15 de agosto de 2018

O estudo de séries temporais tem sido uma área de grande interesse desde a década de 1950 e estudos empíricos vem sendo realizados para avaliar o número cada vez maior de algoritmos disponíveis (LEMKE; GABRYS, 2010).

Entretanto, em mercados financeiros, o sucesso de um investidor depende da qualidade da informação que ele utiliza para auxiliá-lo, e em quão rápido ele consegue chegar a tal decisão (CAVALCANTE et al., 2016).

O problema de seleção de modelos para estudo de séries temporais se assemelha ao problema de seleção de algoritmos encontrados em algumas áreas da computação (PRUDÊNCIO; LUDERMIR, 2004).

Poucos trabalhos foram encontradas com aplicação de meta-aprendizagem em séries temporais financeiras e, nenhum trabalho foi encontrado com uma aplicação direta em dados do mercado de câmbio.

O mercado de câmbio (ou então *forex*), é um dos maiores mercados mundiais e movimenta cerca de dois trilhões de dólares por dia (SHAHBAZI; MEMARZADEH; GRYZ, 2016).

Este é um mercado muito líquido devido ao alto nível de participação, o que oferece uma boa oportunidade de lucros quando os seus ativos apresentam uma tendência forte (NORRIS; BELL; GASKILL, 2009).

Portanto, propomos aplicar um algoritmo de meta-aprendizagem em dados do mercado de câmbio, mais especificamente nos dados do mercado Euro/Dólar, com o intuito de auxiliar investidores na tomada de decisão para a seleção de um modelo.

O presente trabalho traz resultados parciais desenvolvidos durante a pesquisa para obtenção do título de mestre do primeiro autor.

No mercado *forex*, bancos centrais e outras autoridades financeiras são os participantes mais influentes; suas ações afetam tanto os retornos quanto a volatilidade das cotações, visando atingir um valor alvo (CHEN; GAU, 2015).

Baruník, Kočenda e Vácha (2017) apresentam algumas vantagens de se trabalhar com o mercado *forex*:

- 1 operar entre os países 24hrs por dia;
- 2 o volume de negociações diárias é muito superior ao volume de mercados capitais;
- 3 as taxas de câmbio dos pares são mais afetadas por políticas e intervenções monetárias do que as ações e títulos;
- 4 intervenções dos bancos centrais normalmente oferecem um impacto positivo, em especial nos mercados emergentes;

Pincak (2013) relata que modelos lineares são bem utilizados na literatura apesar dos dados financeiros reais apresentarem muitas características não lineares.

Os modelos selecionados são, em sua maioria, modelos lineares e de baixo custo computacional, podendo ser encontrados em bibliotecas como *forecast* do programa R.

- Modelos Autorregressivos Integrados e de Médias Móveis (ARIMA);
- Modelos de Médias Móveis Siples;
- Modelos de Suavização Exponencial;
- Máquinas de Vetores Suporte (SVM);
- Modelos Heterocedásticos (GARCH).

Prudêncio e Ludermir (2004) entendem que a meta-aprendizagem é um processo automático de geração de conhecimento que relaciona a performance dos algoritmos às características do problema.

O Problema de Seleção de Algoritmo (PSA), que foi originalmente proposto por Rice (1976), propõe a existência de uma relação entre características dos problemas e dos algoritmos que podem ser utilizados para solucioná-los.

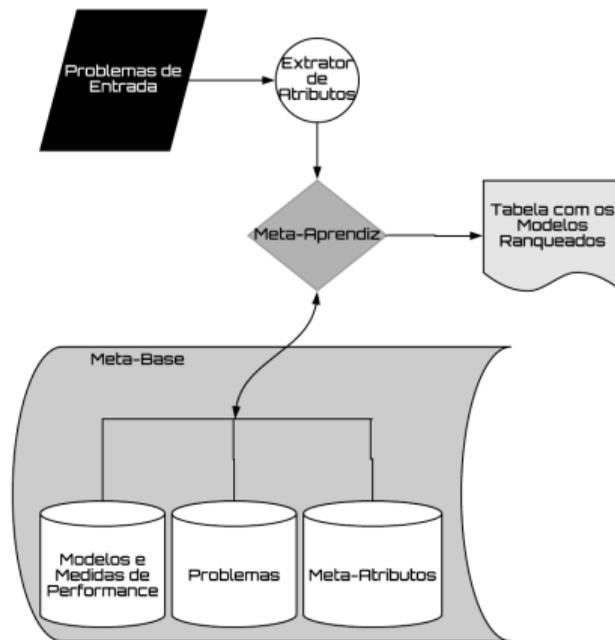


Figura: Fluxograma da Meta-Aprendizagem Utilizada.

Fonte: Autor.

Garcia, Carvalho e Lorena (2016) fornece os principais grupos de características que poderão ser extraídas dos conjuntos de dados:

- Atributos simples, estatísticos e de informação teórica;
- Atributos baseados em modelos;
- Landmarking.

Tripathy, e Panda (2017) apresentam quatro formas de recomendação de algoritmos em seu trabalho, e são elas:

- o melhor algoritmo no conjunto;
- um subconjunto de algoritmos;
- ranqueamento de algoritmos;
- estimativa de performance dos algoritmos.

Tomamos os dados diários do período de 1 de janeiro de 2010 a 31 de dezembro de 2017, por ser um período atual e pela instabilidade econômica apresentada neste período (pós-crise de 2007/2008 e crise do euro nos anos de 2013/2014).

Os dados de treinamento representam, aproximadamente, 75% dos dados coletados e compreende o período de 1 de janeiro de 2010 a 15 de janeiro de 2016, enquanto o restante foram armazenados para testes.

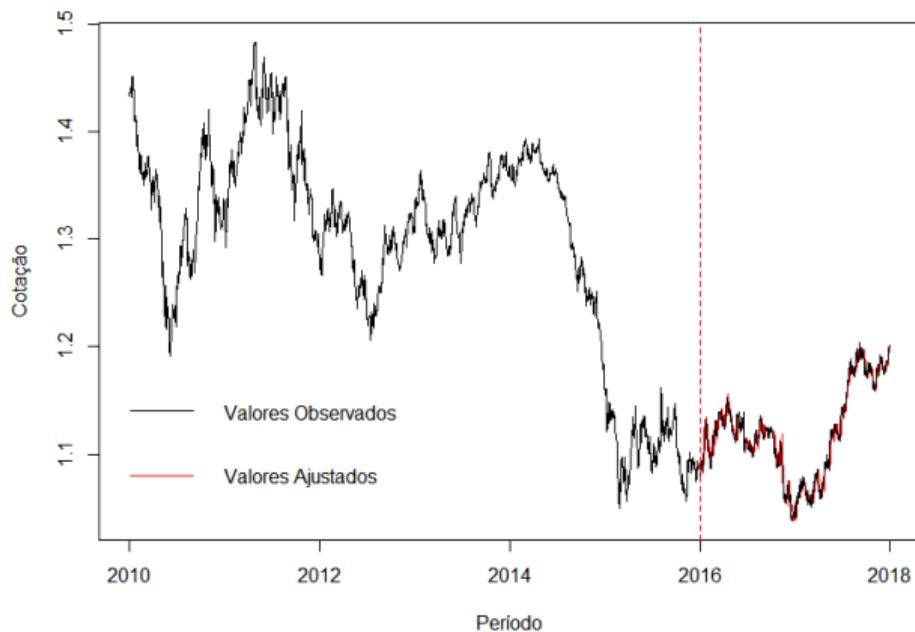


Figura: Gráfico dos dados utilizados.

Fonte: Autor.

A meta-aprendizagem utilizada foi desenvolvida em duas etapas principais: criação da meta-base e teste da meta-aprendizagem.

Para a criação da meta-base, os dados de treinamento foram tomados e separados em sub-conjuntos de 30 dados diários que foram utilizados para estimação dos modelos propostos e extração de atributos.

Os atributos retirados para esta pesquisa foram baseadas nos atributos dados por (LEMKE; GABRYS, 2010) e são resumidos na tabela abaixo.

Tabela: Atributos utilizados na meta-aprendizagem.

<i>Estatísticas gerais</i>
desvio padrão da série assimetria e curtose da série primeiro coeficiente de autocorrelação parcial
<i>medidas centrais</i>
mín., média, média aparada, mediana e máx. intervalo simples e intervalo interquartil pontos de inflexão

Tabela: Número de vitórias, pontuação e média dos ranqueamentos.

	Linear	RBF	ARIMA
Vencedor	244	98	38
Pontuação	1977	1574	1349
Média	3.99	3.17	2.72
	GARCH	MA	Suavização
Vencedor	0	40	76
Pontuação	55	1070	1415
Média	0.11	2.16	2.85

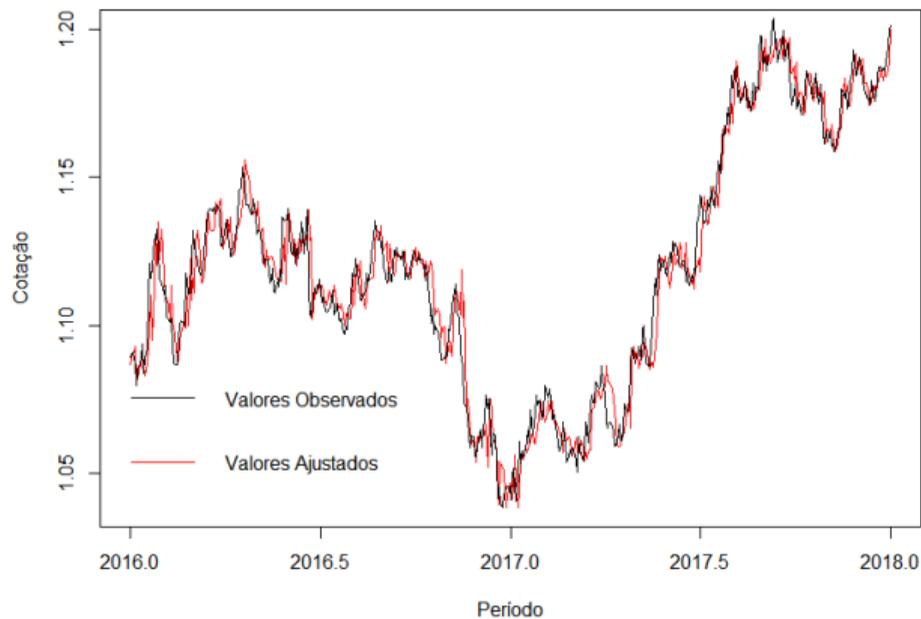


Figura: Gráfico da estimação dada pelas sugestões da meta-aprendizagem.

Fonte: Autor.

Como resultado inicial, a aplicação de meta-aprendizagem demonstrou-se promissora para sugestão de modelos para o estudo de séries temporais. Entretanto, mais alguns testes devem ser realizados para se confirmar a contribuição do algoritmo para o estudo de tal tipo de dado.

Este trabalho demonstra-se motivador ao estender as ferramentas utilizadas para a análise de séries temporais. Para trabalhos futuros, poderá ser acrescentado modelos híbridos na competição de melhor modelo.

 BARUNÍK, J.; KOČENDA, E.; VÁCHA, L. Asymmetric volatility connectedness on the forex market. *Journal of International Money and Finance*, Elsevier BV, v. 77, p. 39–56, oct 2017.

 CAVALCANTE, R. C. et al. Computational intelligence and financial markets: A survey and future directions. *Expert Systems with Applications*, Elsevier BV, v. 55, p. 194–211, aug 2016.

 CHEN, Y.-L.; GAU, Y.-F. Foreign exchange market intervention and price discovery. *Journal of the Japanese and International Economies*, Elsevier BV, v. 38, p. 214–227, dec 2015.

 GARCIA, L. P.; CARVALHO, A. C. de; LORENA, A. C. Noise detection in the meta-learning level. *Neurocomputing*, Elsevier BV, v. 176, p. 14–25, feb 2016.

 LEMKE, C.; GABRYS, B. Meta-learning for time series forecasting and forecast combination. *Neurocomputing*, Elsevier BV, v. 73, n. 10-12, p. 2006–2016, jun 2010.

-  NORRIS, J.; BELL, T.; GASKILL, A. *Mastering the Currency Market: Forex Strategies for High and Low Volatility Markets*. [S.l.]: McGraw-Hill Education, 2009. ISBN 0071634843.
-  PINCAK, R. The string prediction models as invariants of time series in the forex market. *Physica A: Statistical Mechanics and its Applications*, Elsevier BV, v. 392, n. 24, p. 6414–6426, dec 2013.
-  PRUDÊNCIO, R. B.; LUDERMIR, T. B. Meta-learning approaches to selecting time series models. *Neurocomputing*, Elsevier BV, v. 61, p. 121–137, oct 2004.
-  RICE, J. R. The algorithm selection problem. In: *Advances in Computers*. [S.l.]: Elsevier, 1976. p. 65–118.
-  SHAHBAZI, N.; MEMARZADEH, M.; GRYZ, J. Forex market prediction using NARX neural network with bagging. *MATEC Web of Conferences*, EDP Sciences, v. 68, p. 19001, 2016.

 TRIPATHY, M.; ; PANDA, A. A study of algorithm selection in data mining using meta-learning. *Journal of Engineering Science and Technology Review*, Eastern Macedonia and Thrace Institute of Technology, v. 10, n. 2, p. 51–64, jun 2017.