

# UTILIZAÇÃO DE APRENDIZADO DE MÁQUINA PARA CLASSIFICAÇÃO DE TENDÊNCIAS DE RETORNOS DE ATIVOS

Felipe Dias Paiva  
Verônica H. R. da Silva  
Ben-Hur A. Rocha Neto  
Carolina M. Roma  
Gustavo P. Hanaoka



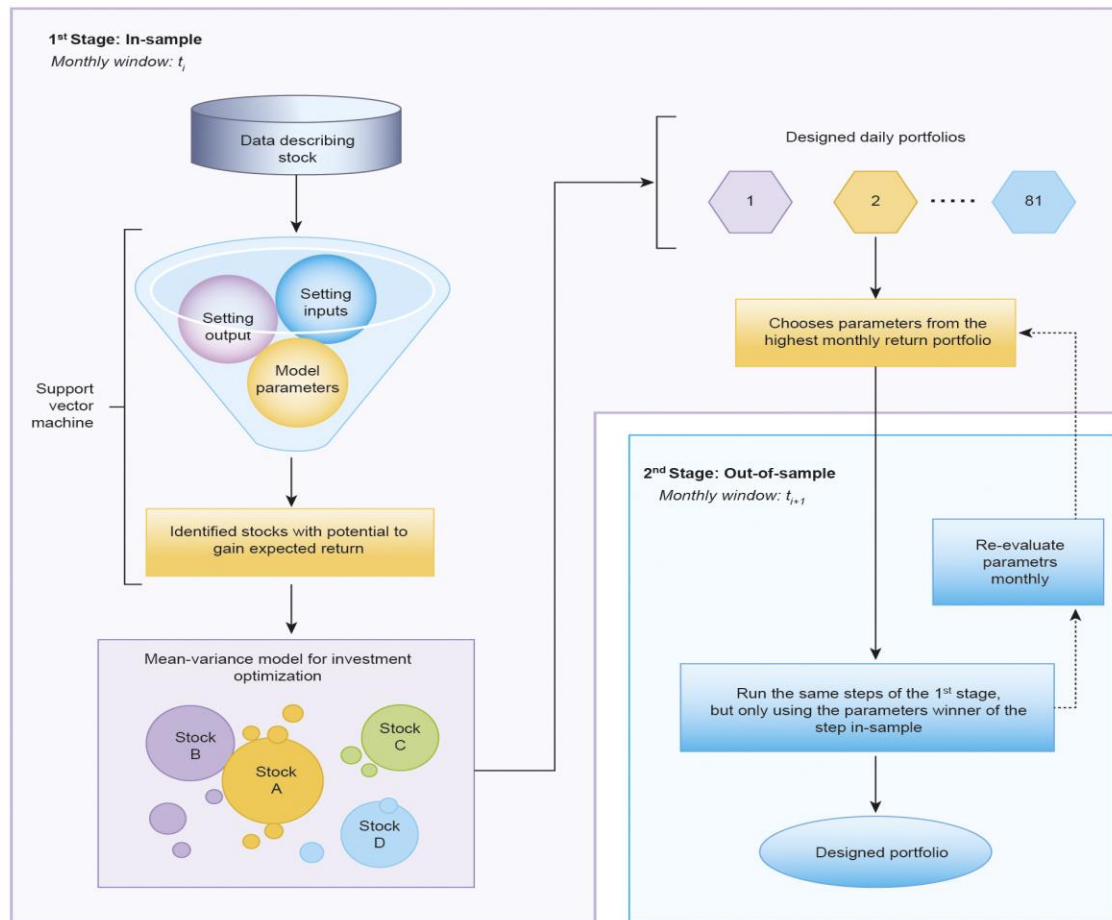
# INTRODUÇÃO

- A previsão dos retornos das ações é considerada uma das tarefas mais desafiadoras do processo que lida com séries temporais financeiras, pois o mercado é essencialmente dinâmico, complexo, evolucionário, não linear, nebuloso, não paramétrico e caótico por natureza.
- Sistemas especialistas têm a sua utilização cada vez mais demandada, como resposta às características do ambiente, buscando identificar umnexo entre passado e futuro, com o objetivo de prever retorno ou o movimento de uma ação.
- As técnicas de inteligência artificial que permitem utilizar dados de diferentes naturezas e ainda não estão sujeitas à rigidez de pressupostos, como os que são impostos aos modelos econométricos.



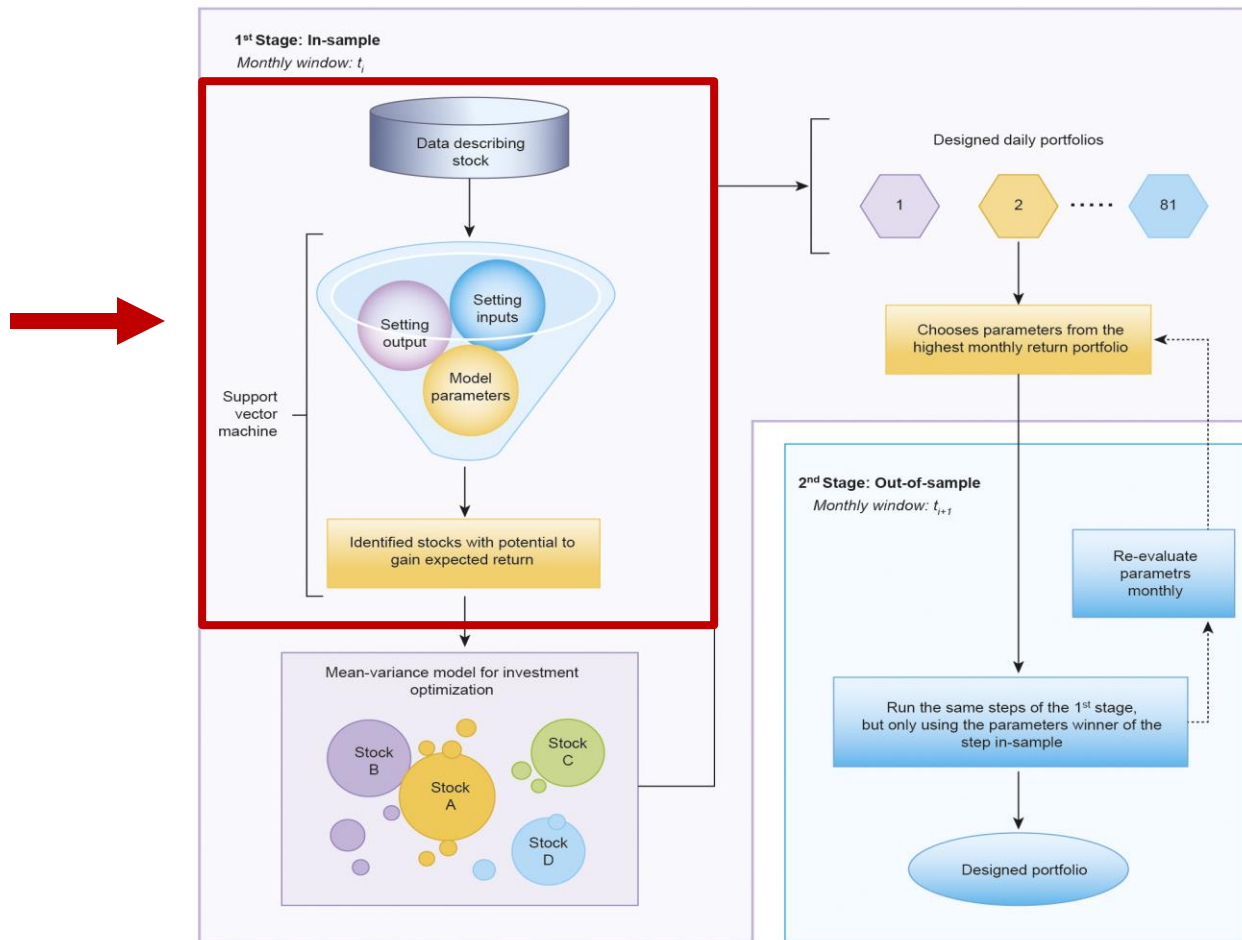
PAIVA, F. D. et al. (2018). Decision-Making for Financial Trading: A Fusion Approach of Machine Learning and Portfolio Selection. *Expert Systems with Applications*.

DOI: 10.1016/j.eswa.2018.08.003



PAIVA, F. D. et al. (2018). Decision-Making for Financial Trading: A Fusion Approach of Machine Learning and Portfolio Selection. *Expert Systems With Applications*.

DOI: 10.1016/j.eswa.2018.08.003



# DESIGN DO EXPERIMENTO

**AMOSTRA:** ativos listados no Ibovespa (135 ativos pesquisados, apenas 19 ativos participaram do Ibovespa ao longo de todo o período).

**PERÍODO:** Janeiro/2002 a Dezembro/2016 (3,716 dias de negociações).

**COLETA DE DADOS:** Terminal Bloomberg.

**DATASET:** Ajustados de Abertura, Máximo, Mínimo, Fechamento e Volume.

**PREVISÃO:** D+1.

**TÉCNICA DE CLASSIFICAÇÃO:** SVM

**TARGETS:** 0.5%; 1.0%; 1.5%; 2.0%.

**TREINAMENTO:** 60 dias, com janela deslizante de 1 dia.

# DESIGN DO EXPERIMENTO

## FEATURES

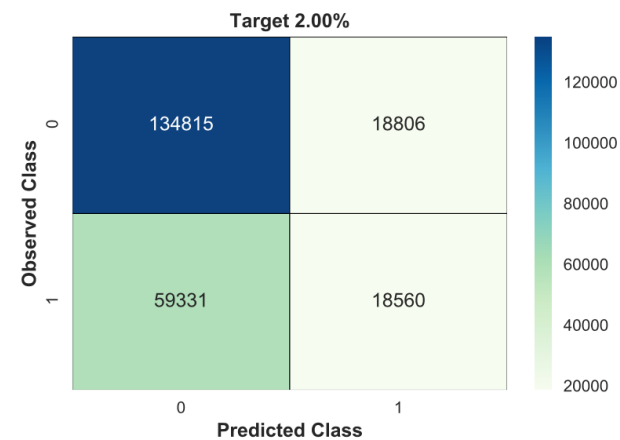
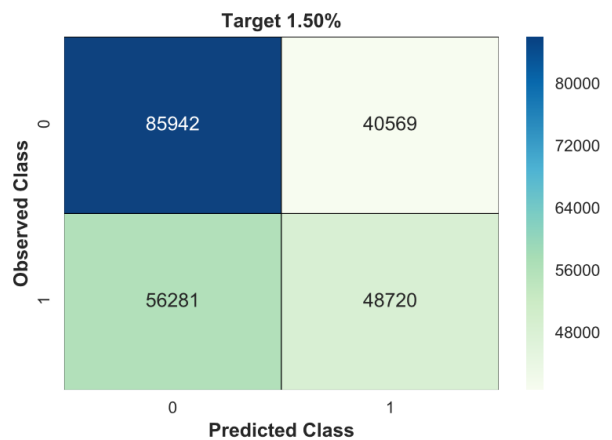
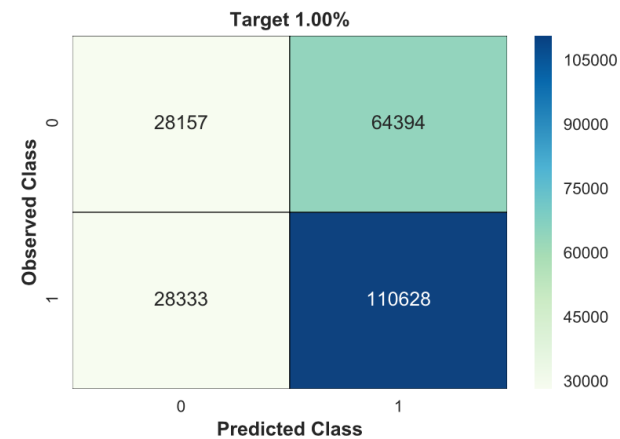
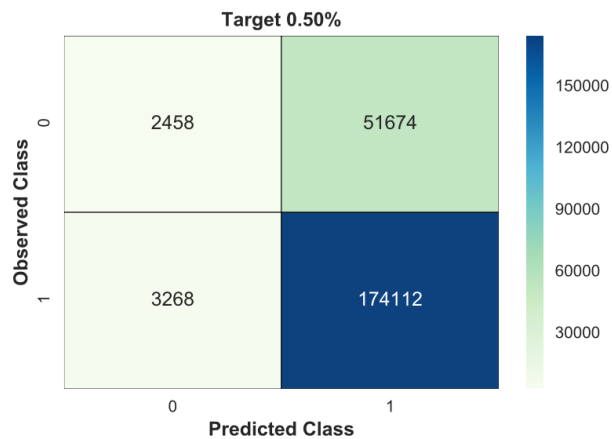
# Atributo	Detalhes	#Atributo	Detalhes
1	$r1 = \ln \left( \frac{\text{close price}_i}{\text{close price}_{i-1}} \right)$	12	$r12 = \ln \left( \frac{\text{low price}_i}{\text{open price}_i} \right)$
2	$r2 = \ln \left( \frac{\text{close price}_{i-1}}{\text{close price}_{i-2}} \right)$	13	$r13 = \ln \left( \frac{\text{low price}_{i-1}}{\text{open price}_{i-1}} \right)$
3	$r3 = \ln \left( \frac{\text{close price}_{i-2}}{\text{close price}_{i-3}} \right)$	14	$r14 = \ln \left( \frac{\text{low price}_{i-2}}{\text{open price}_{i-2}} \right)$
4	$r4 = \ln \left( \frac{\text{close price}_{i-3}}{\text{close price}_{i-4}} \right)$	15	$r15 = \ln \left( \frac{\text{low price}_{i-3}}{\text{open price}_{i-3}} \right)$
5	$r5 = \ln \left( \frac{\text{high price}_i}{\text{open price}_i} \right)$	16	Momentum
6	$r6 = \ln \left( \frac{\text{high price}_i}{\text{open price}_{i-1}} \right)$	17	Relative Strength Index
7	$r7 = \ln \left( \frac{\text{high price}_i}{\text{open price}_{i-2}} \right)$	18	Parabolic SAR
8	$r8 = \ln \left( \frac{\text{high price}_i}{\text{open price}_{i-3}} \right)$	19	Average True Range
9	$r9 = \ln \left( \frac{\text{high price}_{i-1}}{\text{open price}_{i-1}} \right)$	20	True Range
10	$r10 = \ln \left( \frac{\text{high price}_{i-2}}{\text{open price}_{i-2}} \right)$	21	Chaikin A/D Line
11	$r11 = \ln \left( \frac{\text{high price}_{i-3}}{\text{open price}_{i-3}} \right)$	22	On Balance Volume

## MÉTRICAS DE AVALIAÇÃO

Acurácia	Precision	Recall	Specificity	F1 Score
$\frac{TP+TN}{TP+TN+FP+FN}$	$\frac{TP}{TP+FP}$	$\frac{TP}{TP+FN}$	$\frac{TN}{TN+FP}$	$\frac{2*Precision*Recall}{Precision+Recall}$



# RESULTADOS



Total de Oportunidades para Operação: 231.512



# RESULTADOS

## MÉTRICAS DE AVALIAÇÃO

MEDIDA	0.50 %	1.00 %	1.50%	2.00 %
Acurácia	76.27%	59.95%	58.17%	66.25%
Recall	98.16%	76.62%	46.33%	23.83%
Specificity	4.54%	30.42%	67.93%	87.76%
Precision	77.11%	63.21%	54.56%	49.67%
F1 Score	86.37%	70.47%	50.11%	32.21%





# RESULTADOS

## MÉTRICAS DE AVALIAÇÃO

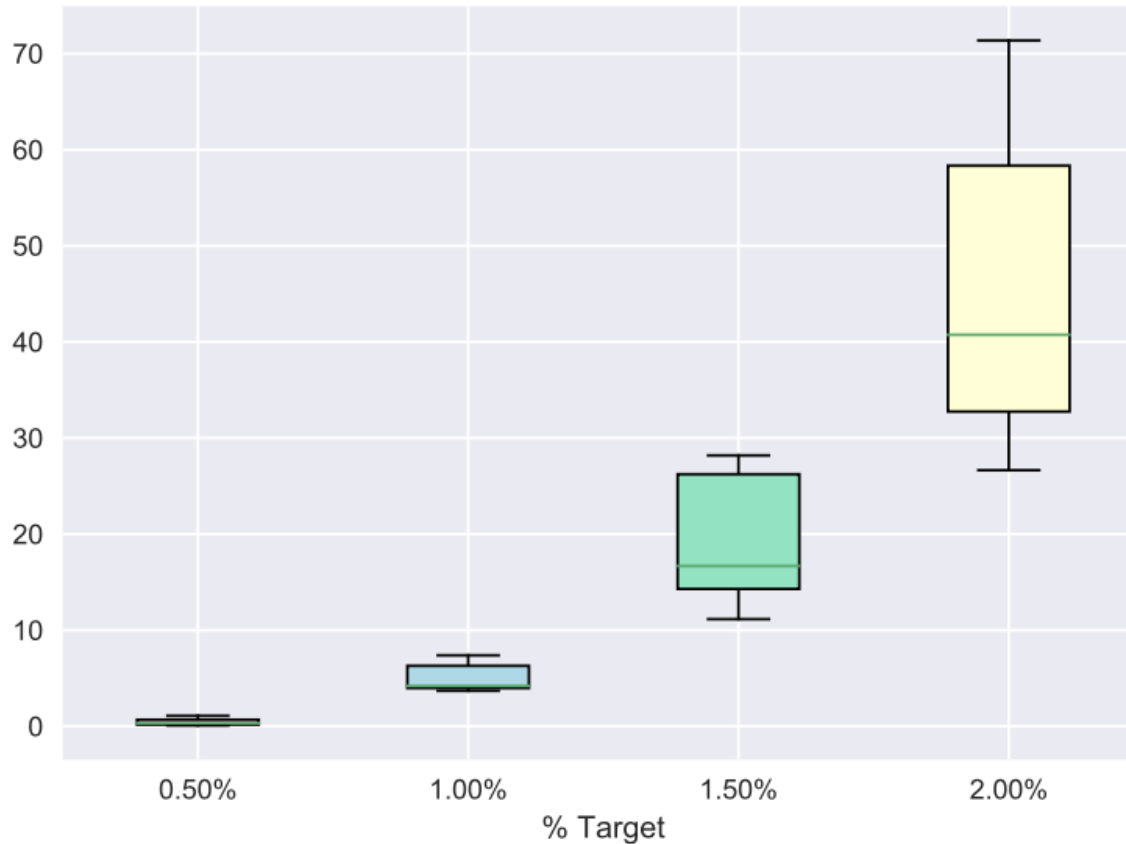
MEDIDA	0.50 %	1.00 %	1.50%	2.00 %
Acurácia	76.27%	59.95%	58.17%	66.25%
Recall	98.16%	76.62%	46.33%	23.83%
Specificity	4.54%	30.42%	67.93%	87.76%
F1 Score	86.37%	70.47%	50.11%	32.21%

## Precision

MEDIDA	0.50 %	1.00 %	1.50%	2.00 %
Naive	76.62%	60.02%	45.35%	33.64
SVM	77.11%	63.21%	54.56%	49.67%
Variação	0.64%	5.31%	20.31%	47.65%

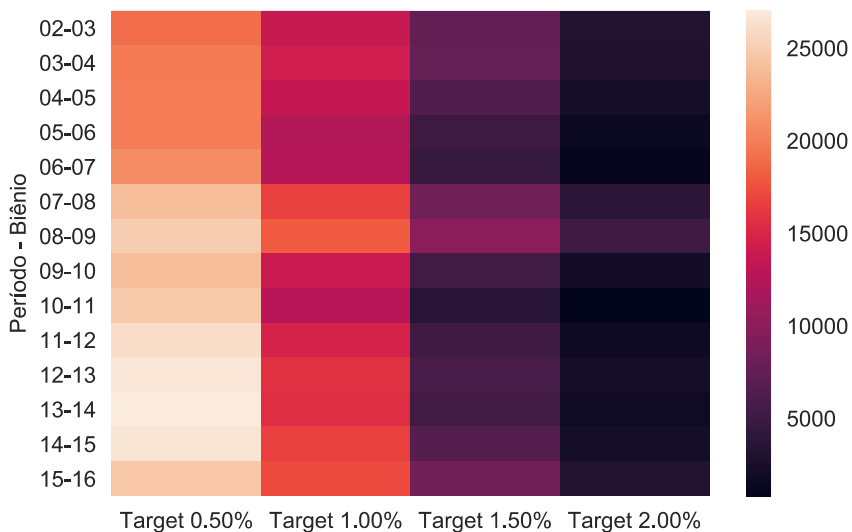
# RESULTADOS

## Diferença entre Precision Naïve e SVM por ativos

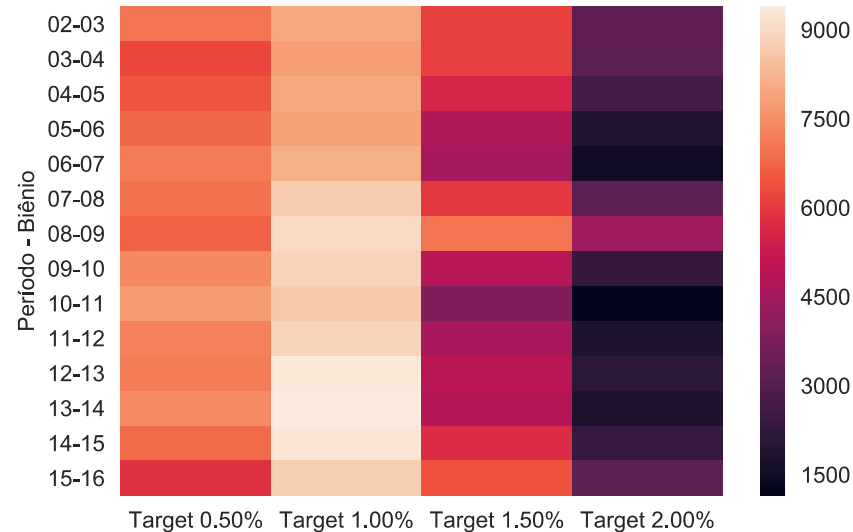


# RESULTADOS

## True Positive

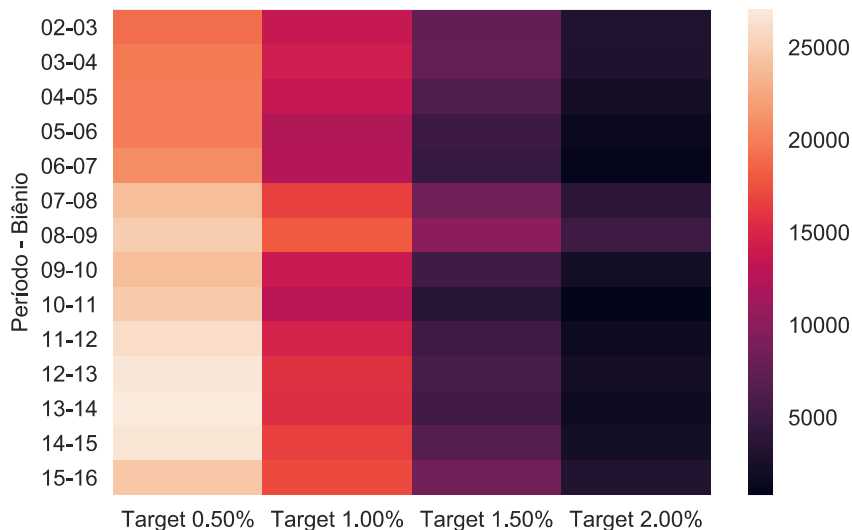


## False Positive

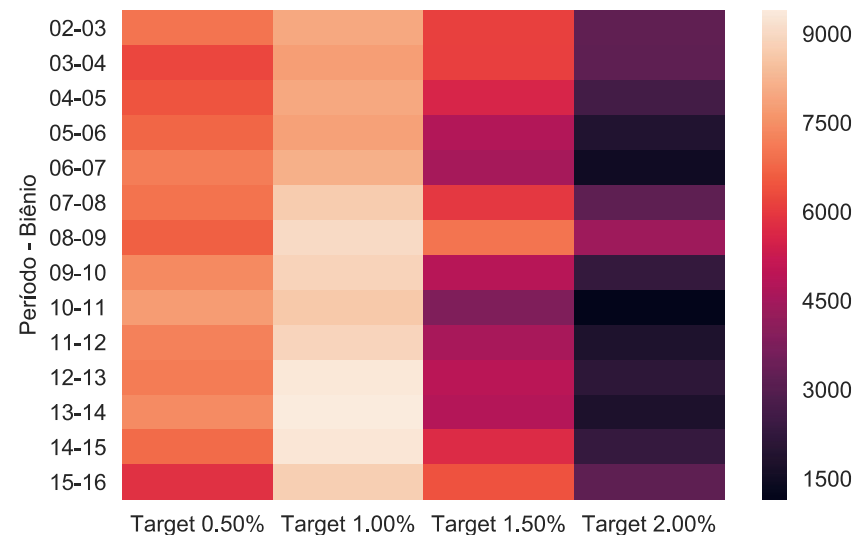


# RESULTADOS

## True Positive



## False Positive



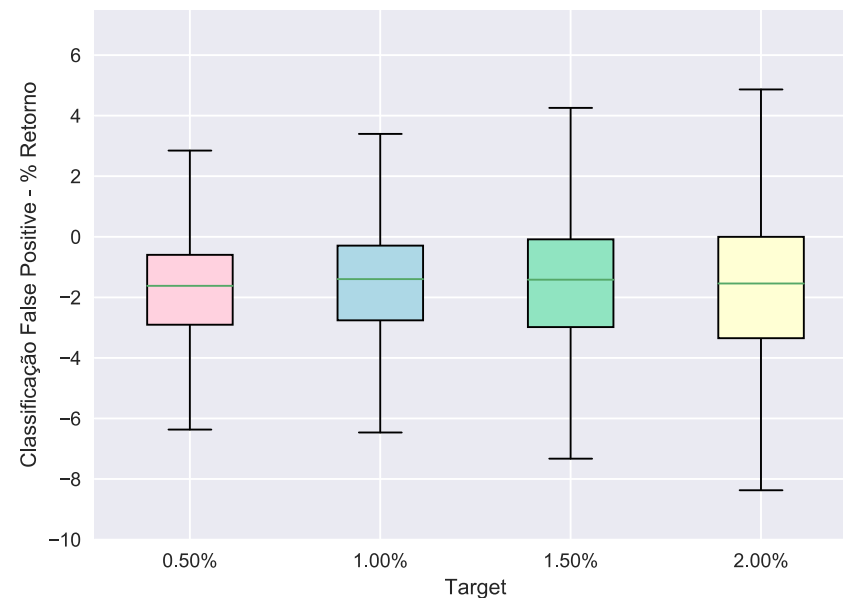
**Ratio médio  
TP/FP**

- 0.5%: 3.40
- 1.0%: 1.71
- 1.5%: 0.95
- 2.0%: 0.95



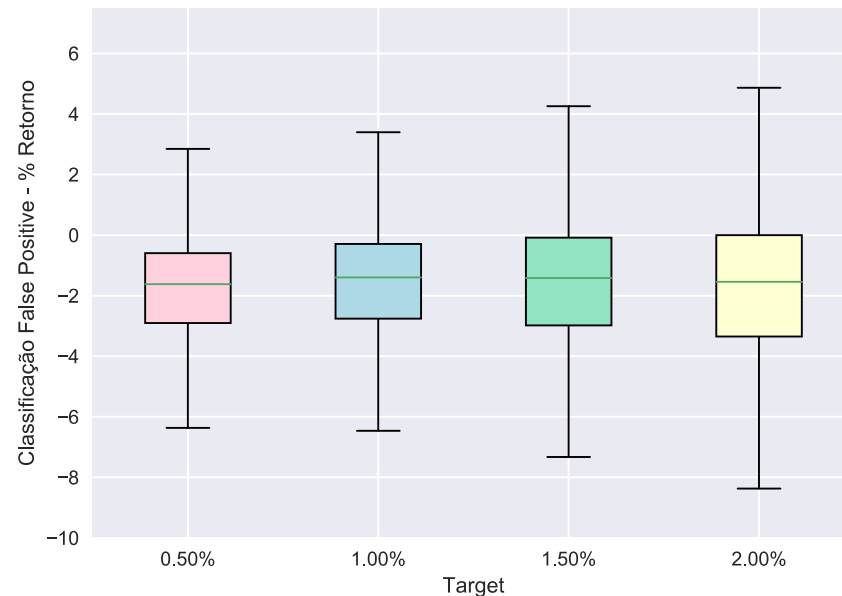
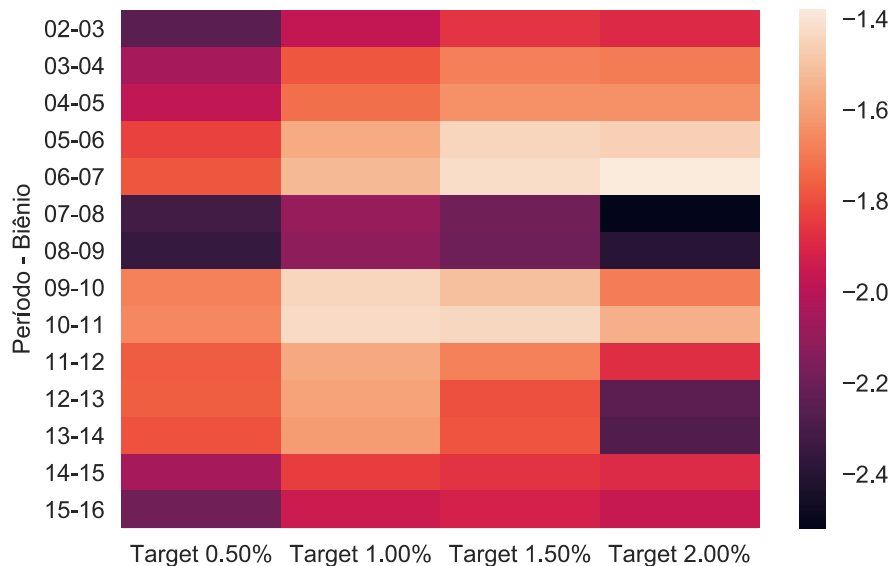
# RESULTADOS

## Perda Média



# RESULTADOS

## Perda Média



**Perda média**

0.5%: 1.97%

1.0%: 1.75%

1.5%: 1.79%

2.0%: 1.98%



# RESULTADOS

**Ratio Médio  
TP/FP**

**0.5%: 3.40**

**1.0%: 1.71**

**1.5%: 0.95**

**2.0%: 0.95**

**X**

**Perda  
Média**

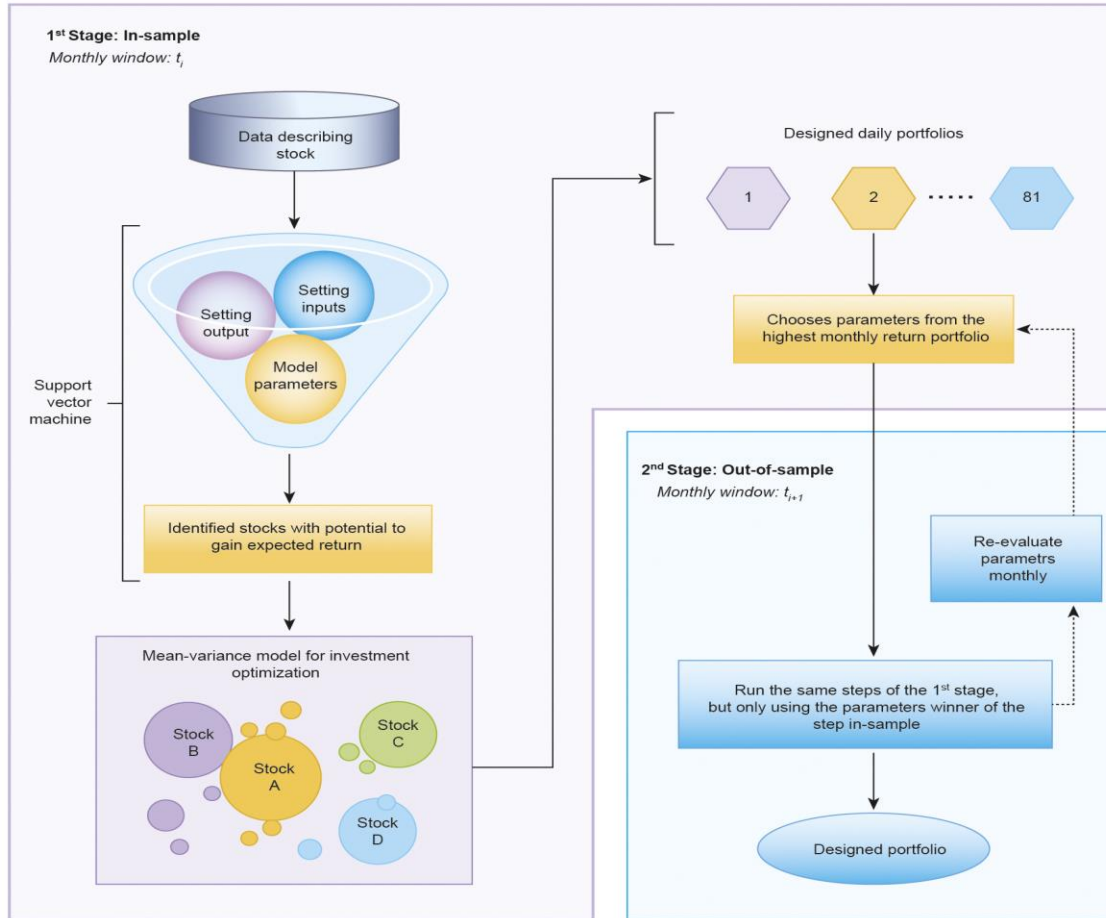
**0.5%: 1.97%**

**1.0%: 1.75%**

**1.5%: 1.79%**

**2.0%: 1.98%**

PAIVA, F. D. et al. (2018). Decision-Making for Financial Trading: A Fusion Approach of Machine Learning and Portfolio Selection. *Expert Systems with Applications*. DOI: 10.1016/j.eswa.2018.08.003





# CONSIDERAÇÕES FINAIS

- Prever quando não operar pode ser mais ou tão importante quanto as indicações de oportunidades de negócio.
- Ratificar a importância da escolha das métricas de avaliação.
- Demonstrado que o desempenho do classificador é sensivelmente impactado com a variação do target.



# TRABALHOS FUTUROS:

- Implementação de outras técnicas de classificação e otimização.
- Calibração do modelo de operação (stop loss).
- Incorporação de análise de sentimento.
- Robot do modelo.



# OBRIGADO!!!

**Felipe Dias Paiva**  
**fpaiva@cefetmg.br**

