

idp

idn

MESTRADO PROFISSIONAL EM ECONOMIA

**OTIMIZAÇÃO DE PORTFÓLIO COM PRÉ-SELEÇÃO DE
ATIVOS USANDO MACHINE LEARNING: UMA APLICAÇÃO NO
CONTEXTO DOS MERCADOS EMERGENTES**

THIAGO RAYMON CRUZ CACIQUE DA COSTA

Brasília-DF, 2022

THIAGO RAYMON CRUZ CACIQUE DA COSTA

OTIMIZAÇÃO DE PORTFÓLIO COM PRÉ-SELEÇÃO DE ATIVOS USANDO MACHINE LEARNING: UMA APLICAÇÃO NO CONTEXTO DOS MERCADOS EMERGENTES

Dissertação apresentada ao Programa de Pós Graduação em Economia, do Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa, como requisito parcial para obtenção do grau de Mestre.

Orientador

Professor Dr. Alexandre Xavier Ywata de Carvalho e Prof. Me. Mathias Schneid Tessmann

Brasília-DF 2022

THIAGO RAYMON CRUZ CACIQUE DA COSTA

OTIMIZAÇÃO DE PORTFÓLIO COM PRÉ-SELEÇÃO DE ATIVOS USANDO MACHINE LEARNING: UMA APLICAÇÃO NO CONTEXTO DOS MERCADOS EMERGENTES

Dissertação apresentada ao Programa de Pós Graduação em Economia, do Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa, como requisito parcial para obtenção do grau de Mestre.

Aprovado em 14 / 02 / 2022

Banca Examinadora

Prof. Dr. Alexandre Xavier Ywata de Carvalho - Orientador

Prof. Dr. Mathias Schneid Tessmann

Prof. Dr. Sergio Jurandyr Machado

Prof. Dr. Rodrigo Soares Holz de Abreu

C837o Costa, Thiago Raymon Cruz Cacique da
Otimização de portfólio com pré-seleção de ativos usando Machine Learning: uma aplicação no contexto dos mercados emergentes / Thiago Raymon Cruz Cacique da Costa. – Brasília: IDP, 2022.

46 p.
Inclui bibliografia.

Trabalho de Conclusão de Curso (Dissertação) – Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa – IDP, Mestrado em Economia, Brasília, 2022.

Orientador: Prof. Dr. Alexandre Xavier Ywata de Carvalho.

1. Otimização de portfólio. 2. Machine Learning. 3. Hierarchical Risk Parity.
4. Mercados emergentes. I. Título.

CDD: 341.38

Ficha catalográfica elaborada pela Biblioteca Ministro Moreira Alves
Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa



RESUMO

A integração da pré-seleção dos ativos na formação de portfólio pode melhorar o desempenho do modelo de otimização de portfólio original (Ma, Han, & Wang, 2021). Uma série de estudos mostra que os modelos de aprendizado de máquina levam superioridade aos métodos de tradicionais de série temporal no aspecto da modelagem da expectativa de retorno. Sendo assim, continuando os estudos de Yu et al. (2020), Wang et al. (2020) e Ma, Han, & Wang (2021), este projeto de pesquisa tem como intuito responder ao apelo de Markowitz (1952, p. 77) para formar melhores crenças sobre o desempenho futuro dos ativos, combinando a pré-seleção de ativos na formação de portfólio por meio da machine learning, aplicando tais técnicas no contexto mercado dos mercados emergentes. Para tanto, será analisado o método tradicional de otimização de carteira de Markowitz (1952), além do Hierarchical Risk Parity, juntamente com três modelos de aprendizagem de máquina: Random Forest, Support Vector Regression e K-Nearest Neighbor, sendo comparado com a estratégia ingênua de diversificação e com os modelo de otimização sem os algoritmos de pré-seleção. Como amostra, serão utilizados 10 ETFs internacionais que reproduzem os índices do mercado acionário de cada mercado emergente analisado do período. Os resultados empíricos mostram que o uso das técnicas de pré-seleção por aprendizagem de máquina melhora o desempenho dos portfólios em relação ao retorno e o prêmio pelo risco, sendo as técnicas que envolvem otimização de portfólio de MV as que melhor performaram. Já em relação à volatilidade, a pré-seleção de ativos por machine learning não resultou em vantagens significativas em nenhuma das medidas estudadas.

Palavras-chave: Otimização de portfólio; Machine learning; Hierarchical Risk Parity; Mercados emergentes.

LISTA DE ABREVIATURAS E SIGLAS

Arima	Autoregressive Integrated Moving Average
Atr	Autoregressive Integrated Moving Average
Cvar	Conditional Value at Risk
Ema	Exponential Moving Average
Ewma	Exponentially Weighted Moving Average
Etf	Exchange Traded Funds
Hrp	Hierarchical Risk Parity
Knn	K-Nearest Neighbor
Md	Maximum Drawdown
Mv	Média Variância
Psar	Parabolic SAR
Obv	On Balance Volume
Omv	Otimização por Média Variância
Rf	Random Forest
Rsi	Relative Strength Index
Svr	Support Vector Regression
Var	Value at Risk

LISTA DE ILUSTRAÇÕES

Figura 1

Retorno acumulado dos portfólios em comparação com o EEM

33

Figura 2

Os 10 maiores drawn downs dos portfólios

34

Figura 3

Volatilidade móvel de 6 meses dos portfólios em relação ao benchmark EEM

36



LISTA DE TABELAS

Tabela 1

Variáveis de Entrada

.....23

Tabela 2

Ativos Utilizados no Estudo

.....26

Tabela 3

Estatística Descritiva dos Retornos em Porcentagem

.....29

Tabela 4

Matriz de Correlação

.....29

Tabela 5

Avaliação de Performance

.....31



SUMÁRIO

1 INTRODUÇÃO	9
2 REFERENCIAL TEÓRICO	12
3 MÉTODO	18
3.1 Random forest (rf)	19
3.2 Support vector regression (svr)	20
3.3 K-nearest neighbor (knn)	20
3.4 Seleção de variável de entrada	21
3.4.1 Relative strength index (rsi)	21
3.4.2 Average true range (atr)	21
3.4.3 On balance volume (obv)	21
3.4.4 Parabolic sar (psar)	21
3.4.5 Exponentially weighted moving average (ewma)	22
Average true range (atr)	23
On balance volume (obv)	23
3.5 Processo de formação do portfólio ótimo	23
3.6 Medidas de desempenho	24
3.6.1 Índice de sharpe	24
3.6.2 Maximum drawdown (md)	24
3.6.3 Conditional value at risk (cvar)	25
3.6.4 Beta	25
3.7 Processo experimental	26
4 RESULTADOS	28
5 CONCLUSÃO	37
Referências Bibliográficas	39



1

INTRODUÇÃO

O marco inicial da teoria moderna de portfólio é o artigo seminal de Markowitz (1952). Em tal estudo, foi introduzido a noção da otimização por programação quadrática para a seleção de uma carteira de investimentos de média variância (MV). Apesar da notória contribuição de Markowitz (1952), o modelo sugerido pelo autor apresenta algumas limitações, levando a montagem de portfólios que não são realmente ótimos (Michaud, 1989).

A aplicação do modelo de MV cria o desafio de retratar o custo de oportunidade entre risco e retorno. Para tanto, o modelo de Markowitz (1952) leva em consideração apenas o primeiro e o segundo momento de retornos históricos para tomada de decisão. Contudo, Yu et al (2020) afirma que a qualidade dos *inputs* e os métodos usados para interpretar as informações afetam a eficácia do métodos de formação de carteira de investimentos. Dessa forma, diversas abordagens foram desenvolvidas para dar prosseguimento ao modelo de MV, com o intuito de entender se apenas a média e a variância histórica dos retornos são suficientes para a otimização de portfólios, além de se entender se uma pré-seleção dos ativos ajuda a criar melhores carteiras.

Estudos como os de Längkvist, Karlsson, & Loutfi (2014) e Baek & Kim (2018) mostram que estatística tradicional, como a utilizada como base no modelo de MV, podem não ser efetivas ao lidar com séries históricas complexas, multidimensionais e ruidosas. Em adição, segundo Perrin & Roncalli (2020, p. 261), uma série de estudos proveram sólidos modelos teóricos para o avanço da discussão de formação de carteira, mas que falham no tratamento de uma grande amostra de ativos e dados para análise. Contudo, o surgimento e o rápido desenvolvimento de algoritmos de aprendizagem de máquina e *big data* traz uma nova luz ao estudo da construção de portfólio, possibilitando ir além da abordagem tradicional de otimização de média-variância.

De acordo com Huang (2012), a linha de pesquisa de otimização de portfólio depende da volatilidade de longo prazo dos dados de séries temporais financeiras no passado, bem como de uma previsão confiável de desempenho de ativos no futuro. O estudo de Guerard,

Markowitz, & Xu (2015) acrescenta que, no processo de gestão de carteira de investimentos, é um fator crítico para a análise do retorno esperado fazer uma seleção preliminar de ativos. Em contrapartida Wang et al (2020) afirma que poucas pesquisas dão atenção à pré-seleção de ativos antes de formar uma carteira.

Uma das bases que fundamentam um modelo de otimização de carteira, assim como introduzido por Markowitz (1952), é a definição da relação que existe entre as movimentações dos retornos dos ativos analisados. Em suma, baixas correlações entre os retornos aumentam a diversificação do portfólio. A pesquisa de Goetzmann, Li, & Rouwenhorst (2005) mostra que a evolução do mercado financeiro resultou em uma maior variedade de investimentos para os investidores em uma escala global. Em contrapartida, como consequência da globalização, a correlação dos principais mercados aumentou. Sendo assim, os benefícios para a diversificação internacional foram recentemente impulsionados pela existência de mercados de capitais emergentes.

Dado o contexto apresentado, dando sequência aos estudos de Yu et al. (2020), Wang et al. (2020), Ma, Han, & Wang (2021), a presente pesquisa tem como objetivo estudar a otimização de portfólio com pré-seleção de ativos usando machine learning no contexto dos mercados emergentes.

O texto está dividido em 3 partes. Além dessa seção introdutória, a segunda parte trará uma explanação introdutória, evidenciando alguns dos estudos-chaves para a construção do conhecimento a teoria de portfólios e machine learning. Por fim, a terceira seção explanará os métodos utilizados, o benchmark considerado e a amostra analisada.



2

REFERENCIAL TEÓRICO

A contribuição de Harry Markowitz para a Teoria Moderna de Portfólio é indiscutível. Além do clássico trabalho de 1952, Markowitz (1956) propôs um algoritmo para resolver problemas de programação quadrática, que teve sua continuidade por diversos outros estudos-chaves como os de Fran & Wolfe (1956), Beale (1959) e Wolfe (1959). Nessas pesquisas de seminários de Markowitz, foi apresentada a fundamentação da otimização por mínima variância (OMV), conforme a Equação 1.

$$\begin{aligned} \min \sigma_p^2 &= \sum_{i=1}^N \sum_{j=1}^N x_i x_j \sigma_{i,j} \\ \text{s. a. } \sum_{i=1}^N x_i &= 1 \end{aligned} \quad (1)$$

em que,

N representa o número de ativos;

σ_p^2 é o desvio padrão do portfólio P

x_i é o peso no ativo i ; e

$\sigma_{i,j}$ Representa a covariância entre dois ativos i e j .

A Equação 1 mostra um processo de minimização do risco de uma carteira P , dado a restrição da soma do peso de todos os ativos que resulta em 1. Sendo assim, a solução de tal programação quadrática provê uma carteira com ponderações tidas como eficientes, minimizando a volatilidade do portfólio, dado uma expectativa de retorno requerido. A Equação 2 define o cálculo da esperança de retorno \bar{r} do portfólio P referida.

$$\bar{r}_p = \sum_{i=1}^N x_i \bar{r}_i \quad (2)$$

Perrin & Roncalli (2020, p. 262) afirma que o sucesso da OMV é devido ao sucesso da programação quadrática. Pela facilidade de resolver tal processo de programação, diversos *softwares* disponibilizam a OMV aos seus usuários de uma forma prática, ajudando o método de média variância a ser usado pela maioria dos gestores de portfólio.

Apesar do brilhantismo da teoria de Harry Markowitz, as soluções provindas da OMV são discutidas em uma série de pesquisas, principalmente focando em seus problemas práticos. De Prado (2016) aponta que os métodos de programação quadrática requerem a inversão de uma matriz de covariância definida positiva, ou seja, todos os valores próprios devem ser positivos. Esta inversão está sujeita a grandes erros quando a matriz de covariância é numericamente mal condicionada. Em outras palavras, à medida que se é adicionado ativos correlacionados, o número da condição aumenta. Em algum ponto, o número da condição é tão alto que erros numéricos tornam a matriz inversa muito instável. Dessa forma, uma pequena mudança em qualquer entrada levará a portfólios muito diferentes.

Dada essa instabilidade de soluções trazidas pela OMV, De Prado (2016) introduz a abordagem *Hierarchical Risk Parity* (HRP). A HRP aplica teoria dos gráficos e técnicas de *machine learning* para construir um portfólio diversificado com base nas informações contidas na matriz de covariância. No entanto, ao contrário da OMV, o HRP não requer a inversão da matriz de covariância.

Como mostra De Prado (2016) e Burggraf (2021), o algoritmo HRP se baseia em três estágios: *tree clustering*, *quase-diagonalization*, *recursive bisection*. A primeira etapa envolve dividir os ativos em diferentes *clusters* usando o algoritmo *Hierarchical Tree Clustering*. Para dois ativos i e j , a matriz de correlação é transformada em uma matriz de distância de correlação D , conforme Equação 3.

$$D(i, j) = \sqrt{0,5 \times (1 - \rho(i, j))} \quad (3)$$

Posteriormente, calcula-se a distância euclidiana entre todas as colunas de forma pareada, com o intuito de encontrar a matriz de distância \bar{D} , descrito na Equação 4.

$$\bar{D}(i, j) = \sqrt{\sum_{k=1}^N (D(k, i) - D(k, j))^2} \quad (4)$$

Dada a matriz de distância \bar{D} , constrói-se um conjunto de *clusters* U usando uma abordagem recursiva. Logo, o primeiro *cluster* (i^*, j^*) é calculado conforme a Equação 5.

$$U[1] = \arg \min_{i,j} \bar{D}(i,j) \quad (5)$$

Após o cálculo do primeiro *cluster*, a matriz \bar{D} é atualizada calculando as distâncias dos outros ativos do cluster $U[1]$ usando *clustering* de ligação única. O intuito combinar recursivamente os ativos no portfólio em *clusters* e atualizar a matriz de distância até que se fique com apenas um único *cluster*. Portanto, para qualquer ativo i fora do *cluster*, a distância para o cluster recém-formado é encontrado seguindo a Equação 6.

$$\bar{D}(i, U[1]) = \min (\bar{D}(i, i^*), \bar{D}(i, j^*)) \quad (6)$$

O segundo estágio reorganiza as linhas e colunas da matriz de covariância, de forma que os maiores valores fiquem ao longo da diagonal. Como bem descrito por De Prado (2016), a etapa *quase-diagonalization* faz com que investimentos semelhantes sejam colocados juntos na matriz de covariância, e investimentos diferentes sejam colocados distantes.

A etapa final *recursive bisection* envolve a atribuição de pesos reais do portfólio aos ativos. O algoritmo explora a propriedade do portfólio de que a alocação de variação inversa é ideal para uma matriz de covariância diagonal.

A literatura de avaliação prática do HRP ainda é bastante escassa (Burggraf, 2021), porém alguns estudos já trazem resultados relevantes. Raffinot (2018) avalia o desempenho do HRP em dados que consistem em setores S&P 500, múltiplos ativos e ações individuais. A análise empírica aponta que carteiras são verdadeiramente diversificadas e alcançam desempenhos ajustados ao risco significativos. Burgaff (2021) aplica tal técnica a dados de criptomoedas, mostrando que o HRP navega melhor pela volatilidade e pelo risco de cauda em comparação com as estratégias tradicionais baseadas em risco. Além disso, o HRP resulta em carteiras mais diversificadas enquanto, o MV constrói portfólios muito concentrados.

As técnicas de aprendizado de máquina, que integram sistemas de inteligência artificial, buscam extrair padrões aprendidos de dados

históricos - em um processo conhecido como treinamento ou aprendizado para posteriormente fazer previsões sobre novos dados (Xiao, Xiao, Lu, & Wang, 2013, p. 99). O início da indústria da *machine learning* teve um início semelhante aos algoritmos de otimização de carteira. Antes de 1990, o aprendizado estatístico se concentrou principalmente em modelos fáceis de resolver do ponto de vista numérico (Perrin & Roncalli, 2020, p. 271). Após essa data, alguns importantes modelos de aprendizagem de máquina começaram a emergir, entre eles se destacam as Redes Neurais (Rosenblatt, 1958; Bishop, 1995; LeCun *et al.*, 1989, Adya & Collopy, 1998), *Support Vector Machines* (Cortes & Vapnik, 1995; Kara, Boyacioglu, & Baykan, 2011), e a regressão Lasso (Tibshirani, 1996).

Uma série de estudos aplicam *machine learning* no escopo das Finanças e da formação de carteiras. Ban, El Karoui, & Lim (2018) adaptou dois métodos de aprendizagem de máquina, *regularization and cross-validation*, para a otimização de portfólios. Os resultados mostraram que as técnicas superaram o *benchmark*. A pesquisa de Gu, Kelly, & Xiu (2020) aplica técnicas de *machine learning* para a mensuração de prêmio de risco de ativos e indica que as técnicas *trees* e Redes Neurais acrescentam poder preditivo em comparação à modelos não lineares.

O avanço do estudo da aprendizagem de máquina no escopo das Finanças trouxe luz à pré-seleção dos ativos por meio da expectativa de retorno. Dessa forma, Yu *et al.* (2020) uniu a pré-seleção dos ativos por meio da projeção do retorno com o modelo de otimização de portfólio clássico de MV e outros desdobramentos como a minimização do *Downside Risk*, *Linearized Value-at-Risk* e *Conditional Value-at-Risk*. Em tal estudo, utilizou-se a *Autoregressive Integrated Moving Average* (ARIMA) para modelar a previsão de retorno de ativos e considerou-se a estratégia ingênua de diversificação igualmente ponderada como *benchmark*. Os resultados mostraram que a superioridade da previsão de carteiras de risco sobre o *benchmark* adotado varia entre vários modelos de carteira, considerando 14 anos de histórico de retornos históricos diários de *Exchange Traded Funds* (ETF) internacionais. O estudo ainda evidencia que o benefício da inclusão da previsão é maior quando o mercado é menos volátil.

Segundo Ma, Han, & Wang (2021), a integração da previsão de retorno por modelos de *machine learning* na formação de carteira pode melhorar o desempenho dos métodos de otimização de portfólio original. Sendo assim, em seu artigo, os autores incorpora os resultados

preditivos de dois modelos de *machine learning*, sendo eles o *Random Forest* (RF) e o *Support Vector Regression* (SVR) e três modelos de *deep learning* (*LSTM neural network*, *deep multilayer perceptron* e *convolutional neural network*) aos métodos de otimização de portfólio MV e Ômega. Utilizando como amostra o retorno histórico de 9 anos das 100 principais ações do mercado chinês, o estudo de Ma, Han, & Wang (2021) mostrou que os modelos MV e ômega com previsão de retorno de RF superaram os outros.

Com escopo semelhante ao da pesquisa de Ma, Han, & Wang (2021), Wang et al. (2020) propõe um método misto que consiste na aplicação de *deep learning*, mais precisamente da *LSTM neural network*, juntamente com o modelo de MV para a formação ótima de portfólio em conjunto com a pré-seleção de ativos, em que dependências de longo prazo de dados de séries temporais financeiras podem ser capturadas. O experimento usa um grande volume de dados de amostra do índice *UK Stock Exchange 100* entre março de 1994 e março de 2019. Os resultados evidenciaram uma superioridade do modelo aos *benchmarks* propostos.



3



3

MÉTODO

Esta seção apresentará os modelos de *machine learning* que serão considerados no estudo, sendo eles o de *Random Forest* (RF) e *Support Vector Regression* (SVR) e K-Nearest Neighbor (KNN). Em adição a esses métodos de aprendizagem de máquina, serão analisados uma estratégia ingênua de diversificação, a $1/N$, no qual a ponderação de todos os ativos é igual, além do modelo clássico de MV, considerando a média dos retornos histórica como a expectativa de retorno.

3.1 RANDOM FOREST (RF)

O RF foi inicialmente proposto por Ho (1995) e consiste em um modelo não paramétrico e não linear que tem como dos seus principais mérito evitar problemas de *overfitting* (Breiman, 2001). O modelo constrói um conjunto de árvores para melhorar a robustez limitada e o desempenho abaixo do ideal das árvores de decisão (Dudoit, Fridlyand, & Speed, 2002).

Assim como descrito por Ballings et al. (2015), o algoritmo *Binary Recursive Partitioning* (BRP) é usado para realizar uma amostra *bootstrap* com o intuito de construir árvores individuais. O BRP cria um subconjunto aleatório de variáveis candidatas e avalia todas as divisões possíveis de tais variáveis. Sendo assim, o melhor subconjunto é usado para criar um particionamento binário. Todo esse processo descrito é repetido recursivamente dentro de cada partição e para quando o tamanho da partição é igual a 1. Os dois principais parâmetros no modelo de RF são o número de árvores no conjunto e o número de variáveis a serem experimentadas em cada divisão. Assim como sugerido por Breiman (2001) e feito por Ballings et al. (2015), o presente estudo usará 500 árvores e o número de variáveis para a raiz quadrada será o número total de preditores.

3.2 SUPPORT VECTOR REGRESSION (SVR)

O SVR é um modelo clássico da aprendizagem de máquina cujo a origem está na teoria de aprendizagem estatística, que é aplicada para determinar como regular a generalização e descobrir o equilíbrio ideal entre a complexidade da estrutura do modelo e o risco empírico (Ma, Han, & Wang, 2021). O referido modelo usa o princípio de nome *Vapnik's Structural Risk Minimization* para resolver os problemas de regressão.

Assim como utilizado por Ma, Han, & Wang (2021), o presente trabalho utilizará a função de base radial como a função kernel do SVR, assim como demonstrado na Equação 3.

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (3)$$

em que,

γ representa a constante da função de base radial.

3.3 K-NEAREST NEIGHBOR (KNN)

O algoritmo KNN é um algoritmo de aprendizado de máquina supervisionado simples que pode ser usado para resolver problemas de classificação e regressão. Como bem define Ma, Han, & Wang (2021), um parâmetro importante para determinar é o K , pois representa o número de vizinhos mais próximos que serão levados em consideração ao pontuar a amostra desconhecida. A amostra desconhecida receberá uma pontuação que corresponde à proporção de positivos das K amostras. Assim como a pesquisa de Ma, Han, & Wang (2021), será validados o parâmetro K testando todos os valores de $K = 5$. Por fim, será usado a árvore k-d (Bentley, 1975).

3.4 SELEÇÃO DE VARIÁVEL DE ENTRADA

A seleção de variáveis de entrada é uma tarefa fundamental de previsão de séries temporais, assim também como para algoritmos de *machine learning*. Como exemplo, Chen & Hao (2018) utilizam a *Exponential Moving Average* (EMA), *Relative Strength Index* (RSI) e *On Balance Volume* (OBV), entre outros, para desenvolver componentes com o intuito de predição dos retornos de ações por meio de aprendizagem de máquina. Com escopo semelhante, Wang et al. (2020) faz uso de indicadores como *Parabolic SAR* (PSAR), *Average True Range* (ATR) e *lags* dos retornos. Como referência desses dois estudos, foi utilizado na presente pesquisa os indicadores técnicos RSI, ATR, OBV, PSAR e, como acréscimo, o *Exponentially Weighted Moving Average* (EWMA), conforme descrito a seguir e resumido na Tabela 1.

3.4.1 RELATIVE STRENGTH INDEX (RSI)

O RSI é um indicador de momento extremamente popular que mede a velocidade e a mudança dos movimentos de preços. A intenção do indicador é dar momentos de compra ou venda de determinado ativo.

3.4.2 AVERAGE TRUE RANGE (ATR)

O ATR é um indicador de análise técnica que reflete a volatilidade do mercado por meio da decomposição de toda a faixa de preço de um ativo por um período.

3.4.3 ON BALANCE VOLUME (OBV)

O OBV é um indicador de momento que mede o volume positivo e negativo fluxo para prever mudanças no preço das ações.

3.4.4 PARABOLIC SAR (PSAR)

O SAR parabólico é usado para determinar a direção n os preços dos ativos sobem ou descem, além de se ajustar à medida que os preços mudarem.

3.4.5 EXPONENTIALLY WEIGHTED MOVING AVERAGE (EWMA)

A EWMA é uma média móvel projetada em uma série histórica para que as observações mais antigas recebam pesos menores. Em outras palavras, os pesos dos dados caem exponencialmente à medida que novas observações vão surgindo. Um parâmetro fundamental para o EWMA é o alpha, α , no qual define a importância das informações mais recentes no cálculo da média móvel. No presente estudo, é utilizado a variável α com o valor igual a 0,2 para ao cálculo da EWMA com os períodos de 5, 10, 22 e 66 dias tanto para o preço quanto para o volume.

Tabela 1: Variáveis de Entrada

Indicador Técnico	Equação
Relative Strength Index (RSI)	$RSI(n) = 100 - \frac{100}{1 + \left(\frac{EMA(n)_{up}}{EMA(n)_{down}} \right)}$ <p>$EMA(n)_{up}$ é a alta nos preços no período n e $EMA(n)_{down}$ representa a queda.</p>
Average True Range (ATR)	$TR = MAX[(High - Low), Abs(High - Close), Abs(L - Close)]$ $ATR = \left(\frac{1}{n} \right) \sum_{i=1}^n TR_i$
On Balance Volume (OBV)	$OBV(n) = OBV_{n-1} + \begin{cases} volume, & se\ close_n > close_{n-1} \\ 0, & se\ close_n = close_{n-1} \\ volume, & se\ close_n < close_{n-1} \end{cases}$
Exponentially Weighted Moving Average (EWMA)	$EWMA = \alpha \times r_n + (1 - \alpha) \times EWMA_{t-1}$
Parabolic SAR (PSAR)	$PSAR = PSAR_{n-1} + \alpha(EP - PSAR_{n-1})$

3.5 PROCESSO DE FORMAÇÃO DO PORTFÓLIO ÓTIMO

Assim como Fischer & Krauss (2018) e Wang et al. (2020), é considerada cada conjunto de treinamento-teste envolvendo um período de treinamento de 750 dias e um período de teste de 250 dias. Os dados da amostra compreendem o período de setembro de 2009 a setembro de 2021. Dessa forma, os dados dos primeiros 750 dias são usados para treinamento com janelas móveis, e os outros 250 dados restantes são utilizados para testes *out-of-sample*. Em seguida, toda a rede avançará 250 dias e o processo irá se repetir.

Ainda baseado no estudo de Wang et al. (2020), uma vez que todos os ativos estejam previstos, é feita a ordenação de todas os ETFs para o próximo período em ordem decrescente desse retorno esperado. Apenas o top k do ranking com os ativos de maior retorno que são considerados qualificados para entrar na próxima fase. O objetivo da segunda etapa é obter a alocação de capital para cada ativo, logo é onde os modelos MV e HRP são utilizados. Vale esclarecer que o modelo proposto não leva em consideração as preferências de risco dos investidores e ativos livres de risco, portanto, as carteiras são compostas

exclusivamente por ativos de risco. Tendo isso em mente, os recursos disponíveis serão destinados à carteira com a menor variância.

Confirmados os ativos e as respectivas proporções de investimento, o próximo passo é a alocação de capital na abertura do próximo pregão. São comprados os ativos selecionados pelos otimizadores de portfólios *a posteriori* da construção do top k no início do pregão seguinte ao processo *in-sample*. Paiva et al. (2019) mostra que a carteira com sete ativos tem um desempenho melhor do que outras com diferentes números de ativos. Dessa forma, encara-se $k = 7$ no presente estudo.

3.6 MEDIDAS DE DESEMPENHO

Como base no estudo de Bessler, Opfer, & Wolff (2017), a presente pesquisa apresenta várias medidas de desempenho para avaliar as carteiras otimizadas desenvolvidas por cada estratégia. Tais medidas são o Índice de Sharpe, *Maximum Drawdown* (MD), *Conditional Value at Risk* (CVaR) e o Beta.

3.6.1 ÍNDICE DE SHARPE

O Índice de Sharpe é definido como a fração do excesso de retorno (retorno médio, \bar{r} , menos a taxa de juros livre de risco, r_f) *out-of-sample* dado pela estratégia s dividido pelo desvio padrão *out-of-sample* dos retornos, $\hat{\sigma}_s$ assim como definido na Equação 5.

$$IS = \frac{\bar{r}_s - r_f}{\hat{\sigma}_s} \quad (5)$$

3.6.2 MAXIMUM DRAWDOWN (MD)

Conforme proposto por Grossman e Zhou (1993) como medidas alternativas de risco, calcula-se o MD, que é um indicador de risco que evidencia a maior perda ocorrida desde um ponto de alta até um ponto de mínima em uma série histórica de cada estratégia s .

$$MD_s = \text{Max}_{s,t^* \in (0,T)} \left[\text{Max}_{s,t \in (0,t^*)} \left(\frac{P_{s,t} - P_{s,t^*}}{P_{s,t}} \right) \right] \quad (6)$$

em que,

$P_{s,t}$ é o preço do portfólio gerado pela estratégia s no tempo t ; e

P_{s,t^*} denota o preço do portfólio gerado pela estratégia s no tempo t^* , quando o portfólio é vendido.

3.6.3 CONDITIONAL VALUE AT RISK (CVAR)

O *Value at Risk* (VaR) é uma técnica estatística usada para medir o nível de risco financeiro dentro de uma carteira de investimento em um período de tempo específico. Em complemento, CVaR tenta abordar as deficiências do modelo VaR. Enquanto VaR representa uma perda de pior caso associada a uma probabilidade e um horizonte de tempo, CVaR é a perda esperada se esse limite de pior caso alguma vez for ultrapassado, assim como mostrado na Equação 7.

$$CVaR = \frac{1}{1-c} \int_{-1}^{VaR} xp(x)dx \quad (7)$$

onde,

$p(x)dx$ é a densidade de probabilidade de obter um retorno com valor x ;

c denota o ponto de corte da distribuição;

3.6.4 BETA

O coeficiente beta, β , mede a volatilidade dos retornos do portfólio i em comparação com o risco sistemático do mercado.

$$\beta = \frac{\text{Covariância}(R_P, R_M)}{\text{Variância}(R_M)} \quad (8)$$

no qual,

R_P é o retorno de um portfólio individual P ; e

R_M representa o retorno da carteira de mercado;

O cálculo do beta é usado para entender se determinado portfólio se move na mesma direção que uma de mercado. Dessa forma, fornece informações sobre o quão tal portfólio ação é em relação ao resto do mercado. Para o presente estudo, usa-se o como *proxy* de carteira de mercado o EEM, ETF que replica uma carteira teórica das principais ações negociadas no mercado de capitais emergente.

3.7 PROCESSO EXPERIMENTAL

Para verificar a eficácia dos modelos propostos, foram coletadas 10 ETFs internacionais que reproduzem os índices do mercado acionário de cada mercado emergente analisado do período de 1 de setembro de 2009 a 1 de setembro de 2021 para a realização dos testes empíricos, conforme demonstrado na Tabela 2.

Tabela 2: Ativos Utilizados no Estudo.

Ativo	Símbolo	País
Índice iShares MSCI Brazil	EWZ	Brasil
Índice iShares MSCI Chile	ECH	Chile
Índice iShares MSCI Hong Kong	EWH	China (Hong Kong)
Índice iShares MSCI Malaysia	EWM	Malásia
Índice iShares MSCI Mexico	EWX	México
Índice iShares MSCI Singapore	EWS	Singapura



Índice iShares MSCI South Africa	EZA	África do Sul
Índice iShares MSCI South Korea	EWY	Coreia do Sul
Índice iShares MSCI Taiwan	EWT	Taiwan
Índice iShares MSCI Turkey	TUR	Turquia





4

4

RESULTADO

A Tabela 3 mostra algumas estatísticas dos retornos diários dos ETFs estudados e a Tabela 4 traz a matriz de correlação entre tais ativos. É observado pela análise dessas tabelas que a correlação entre os ativos é alta, sendo o menor valor do coeficiente de Pearson de 0,481 entre os ETF's EWH, proveniente de Hong Kong, e TUR, que segue o mercado acionário turco. Os valores altos de correlação comprometem o poder de diversificação independente da técnica utilizada. Além disso, alguns ativos, conforme mostrado na Tabela 3, tem o retorno diário médio negativo, assim como o EWZ, ECH e TUR.

Tabela 3: Estatística Descritiva dos Retornos em Porcentagem.

ETF	Média	Desvio Padrão	Mínimo	Percentil 25%	Percentil 50%	Percentil 75%	Máximo
EWZ	-0,003%	2,204%	-26,257%	-1,151%	0,064%	1,177%	16,232%
ECH	-0,007%	1,500%	-16,972%	-0,768%	0,022%	0,763%	10,179%
EWH	0,032%	1,196%	-9,831%	-0,555%	0,066%	0,663%	6,349%
EWM	0,012%	1,158%	-11,275%	-0,536%	0,034%	0,608%	7,115%
EWV	0,014%	1,570%	-16,547%	-0,763%	0,038%	0,891%	7,915%
EWS	0,019%	1,190%	-10,271%	-0,547%	0,053%	0,643%	7,313%
EZA	0,015%	1,974%	-16,044%	-1,064%	0,090%	1,203%	9,613%
EWY	0,030%	1,564%	-17,204%	-0,748%	0,077%	0,913%	11,730%
EWT	0,046%	1,302%	-11,514%	-0,642%	0,080%	0,768%	6,525%
TUR	-0,013%	2,115%	-21,001%	-1,062%	0,068%	1,225%	10,683%

Tabela 4: Matriz de Correlação.

Correlação de Pearson	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
EWZ (1)	1,000	0,635	0,582	0,640	0,714	0,640	0,702	0,650	0,608	0,501
ECH (2)		1,000	0,563	0,607	0,665	0,611	0,661	0,623	0,598	0,492

EWH (3)			1,000	0,638	0,618	0,754	0,661	0,737	0,717	0,481
EWM (4)				1,000	0,667	0,724	0,708	0,713	0,687	0,520
EWV(5)					1,000	0,698	0,753	0,704	0,650	0,563
EWS (6)						1,000	0,737	0,794	0,773	0,546
EZA (7)							1,000	0,740	0,713	0,597
EWY (8)								1,000	0,819	0,543
EWT (9)									1,000	0,509
TUR (10)										1,000

A Tabela 5 revela o desempenho das técnicas de aprendizagem de máquina em conjunto com os modelos de otimização de portfólio, assim como possibilita a comparação com tais técnicas de montagem de carteira sem a pré-seleção dos ativos por *machine learning*, além da análise da estratégia ingênua 1/N.

Vê-se pela Tabela 5 que o número médio de ativos no portfólio otimizado, como um indicador para a magnitude da diversificação, é maior para os modelos que utilizaram a HRP do que para aqueles que usaram o MV. Consequentemente, o HRP oferece mais diversificação entre os ativos e têm alocações menos extremas. Contudo, é interessante observar que tal diversificação não resultou em um Beta e uma volatilidade menor, assim como seria esperado. Em contrapartida, é possível analisar que os modelos que utilizam MV, mesmo com menos ativos no portfólio, conseguem trazer carteiras com um risco menor, mostrando que a quantidade de ativos mais alta não necessariamente resulta em uma menor volatilidade dos retornos.

Tabela 5: Avaliação de Performance.

Variável	KNN + MV	KNN + HRP	RF + MV	RF + HRP	SVR + MV	SVR + HRP	MV	HRP	1/N
Retorno Anualizado	3,280%	-1,200%	2,770%	-0,102%	0,938%	-2,562%	0,809%	-1,837%	-1,285%
Volatilidade Anualizada	18,182%	19,699%	17,875%	19,276%	17,367%	19,150%	17,310%	18,823%	18,094%
Índice de Sharpe	0,269	0,038	0,243	0,092	0,141	-0,039	0,134	-0,003	0,020
Drawdown Máximo	-40,474%	-49,658%	-40,067%	-44,332%	-39,688%	-48,898%	-39,536%	-48,455%	-46,593%
CVaR	-2,781%	-3,052%	-2,711%	-2,997%	-2,612%	-2,942%	-2,600%	-2,885%	-2,788%
Beta	0,750	0,800	0,687	0,780	0,658	0,720	0,649	0,763	0,724
Retorno Acumulado	37,741%	-11,284%	31,138%	-1,010%	9,703%	-1,427%	8,319%	-16,800%	-12,041%
Número médio de ativos no portfólio	3	7	3,4	7	3,3	7	3	7	10

A Figura 1 mostra o retorno acumulado de cada modelo em comparação com o *benchmark* EEM, que busca acompanhar os resultados dos investimentos de um índice composto por ações de mercados emergentes de grande e média capitalização. Com o auxílio da Tabela 5, pode-se analisar que os modelos MV conseguem superar o *benchmark*, independentemente se estiver acompanhado pela pré-seleção ou não.

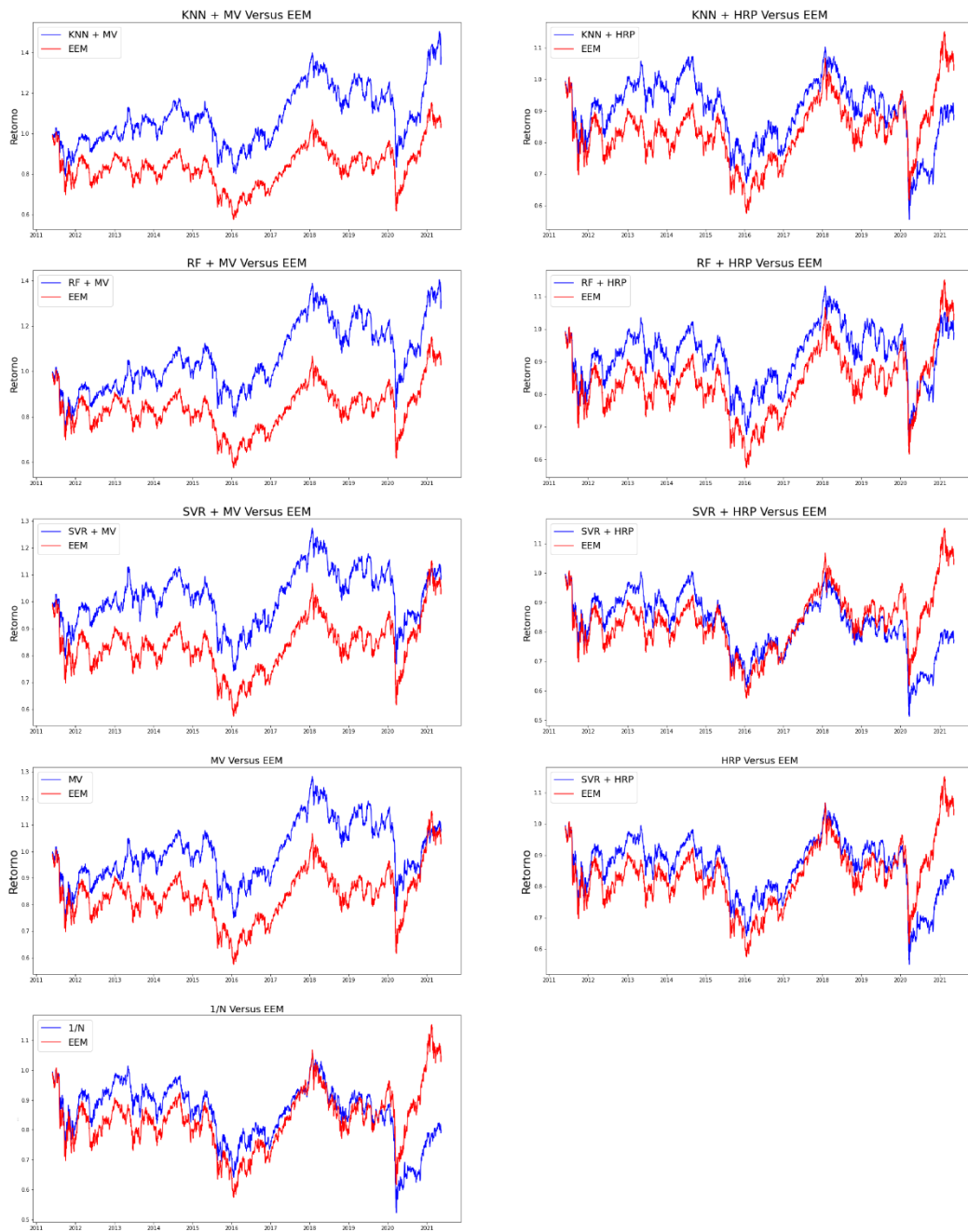
Ainda no escopo do retorno acumulado, é notável que a presença dos modelos de pré-seleção melhora o retorno das estratégias. Em adição, é possível destacar o desempenho da técnica de *machine learning* KNN. Em todas as combinações, a referida técnica provém um ganho significativo quando comparado com o modelo de otimização sem pré-seleção. Como destaque, tem-se o KNN + MV com 37,741% de retorno acumulado contra 8,319% do modelo MV sem pré-seleção. Vê-se também o prejuízo do modelo HRP sendo suavizado quando combinado com o KNN, sendo o KNN + HRP tendo um retorno de -11,284% contra -16,800% sem o uso da aprendizagem de máquina.

A Tabela 5 e a Figura 1 mostram que as carteiras construídas por RF tiveram um retorno acumulado de 31,138%, cujo desempenho foi pior do que as de KNN, mas ainda melhores em comparação aos portfólios 9,703% provenientes da técnica SVR, quando se trata do retorno da combinação com a MV. Já em relação ao processo de otimização de HRP, a melhor técnica de *machine learning* para pré-seleção de dados foi

Em relação ao *benchmark* EEM, a Figura 1 evidencia que apenas os modelos de MV conseguiram superar tal índice. Todos os modelos que se baseiam no HRP são superados pelos retornos do EEM, mesmo auxiliados pela pré-seleção. Quando comparado com a estratégia ingênua 1/N, apenas a otimização por HRP sem uso de aprendizagem de máquina tem retorno acumulado menor.

Em relação ao prêmio pelo risco, a Tabela 5 mostra que o destaque positivo fica para a KNN + MV, com o Índice de Sharpe de 0,269, juntamente com a RF + MV, com 0,243 de valor para a medida de recompensa pelo risco idiossincrático. Em se tratando do Beta, nota-se que o uso da pré-seleção por *machine learning* aumentou o valor de tal índice, quando se comparado aos modelos puros. Como exemplo, a técnica de maior Beta é a KNN + HRP com o valor de 0,800, em comparação com a técnica HRP pura, com o Beta de 0,763.

Figura 1: Retorno Acumulado dos Portfólios em Comparação com o EEM.



A Figura 2 destaca os 10 maiores *draw downs* de cada uma das estratégias no período considerado. As técnicas que consideram MV se destacam novamente com os *draw downs* mais curtos, sendo os principais métodos o KNN + MV e o RF + MV. Com o auxílio da Tabela 4, percebe-se que o modelo de menor queda máxima foi o MV sem o auxílio de *machine learning*, contudo a combinação com os algoritmos de pré-seleção dos ativos mantiveram tal técnica em um nível muito semelhante de *draw down*. Em relação ao HRP, vê-se as maiores quedas máximas, tendo como destaque negativo o KNN + HRP, com

draw down máximo de -49,658%, e o SVR + HRP com a maior queda de -48,898%, sendo o HRP sem pré-seleção de ativos com -48,455%. Dessa forma, nota-se que o uso desses modelos de pré-seleção não provem quedas máximas mais contidas quando considerada a presente amostra. Por fim, a Figura 2 mostra que os *draw downs* provenientes do modelos que envolvem o uso do HRP são mais longos quando comparado aos outros.

Figura 2: Os 10 Maiores *drawn downs* dos Portfólios.

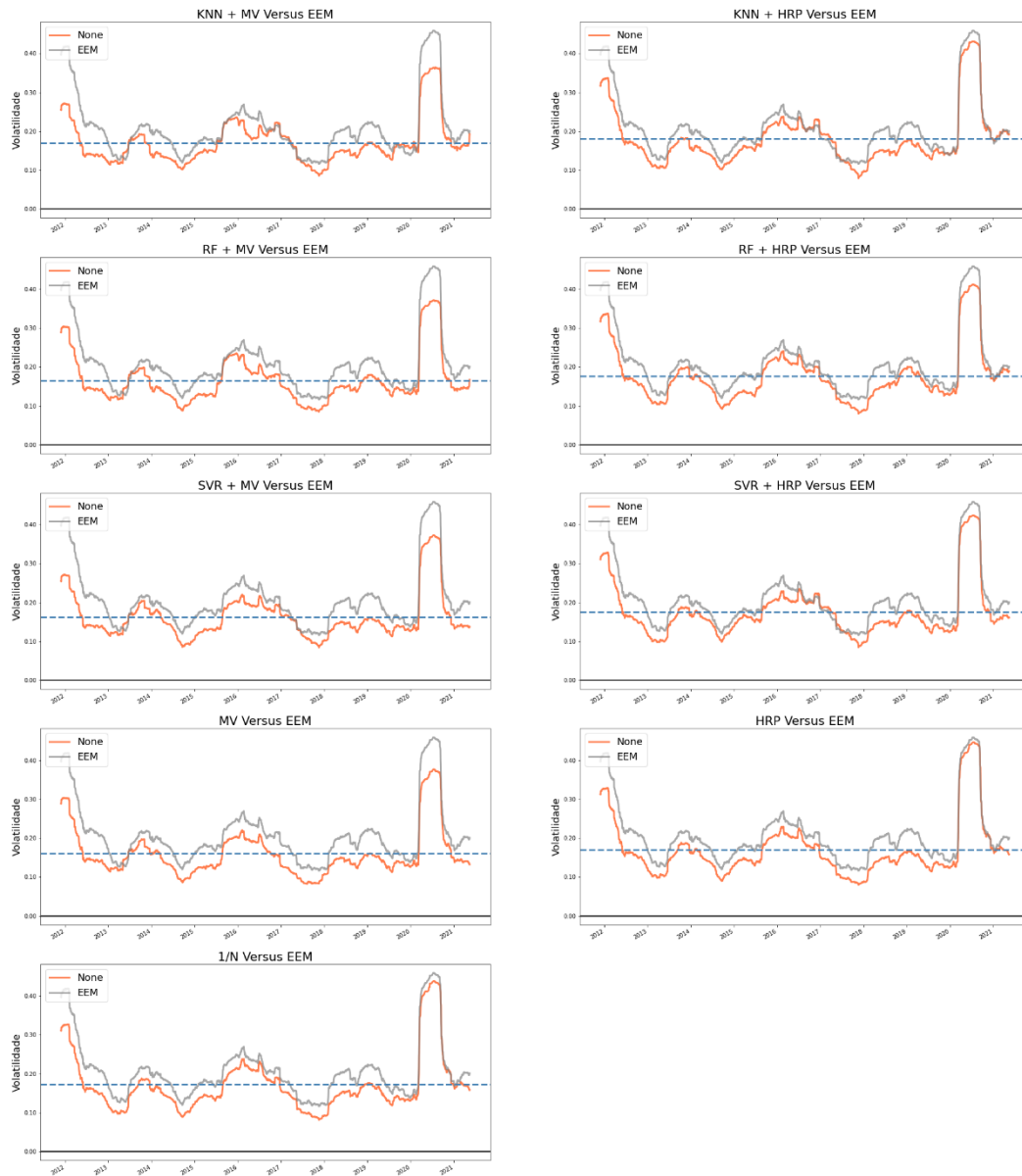


No tocante à análise da volatilidade, a Tabela 4 mostra que os as técnicas de pré-seleção dos ativos em combinação com os modelos de otimização de portfólio proveram portfólios mais agressivos. Em adição, além do Beta e da volatilidade, a medida CVaR é maior quando se utiliza os algoritmos de *machine learning*. Dessa forma, o uso das técnicas de

aprendizagem de máquina na amostra analisada não resultou em uma diminuição das medidas risco.

Ainda no escopo do estudo da variação dos retornos, a Figura 3 mostra uma comparação da volatilidade calculada em uma janela de 6 meses das carteiras provenientes dos modelos propostos em relação ao *benchmark* EEM. Analisando tal figura, vê-se que a volatilidade dos portfólios fica muito próxima ao risco do EEM. Contudo, algumas técnicas se mostraram com picos de volatilidade mais controlados do que outras. Em 2020, ano marcado pelo maior agrupamento de volatilidade da série estudada, os modelos que se utilizam de MV se destacaram positivamente, independente se houve uso da pré-seleção por *machine learning*. Em relação especificamente a esse pico de risco de 2020, os modelos de HRP são os destaques negativos. As técnicas de aprendizagem de máquina conseguiram diminuir o impacto, mas não de forma significativa.

Figura 3: Volatilidade Móvel de 6 meses dos Porfólios em Relação ao Benchmark EEM.





5

5

CONCLUSÃO

A presente pesquisa tem como objetivo estudar a otimização de portfólio com pré-seleção de ativos usando *machine learning* no contexto dos mercados emergentes, com o objetivo de dar continuidade ao trabalho de Yu et al. (2020), Wang et al. (2020), Ma, Han, & Wang (2021). Para tanto, o clássico modelo de Média Variância (MV) de Markowitz (1952) e o *Hierarchical Risk Parity* (HRP) foram combinados com os algoritmos de aprendizagem de máquina *Random Forest* (RF), *Support Vector Regression* (SVR) e *K-Nearest Neighbor* (KNN) para a pré-seleção dos ativos. O desempenho das técnicas foi analisado levando em conta o retorno anualizado, risco anualizado, *drawdown* máximo, Índice de Sharpe, CVaR, Beta e o retorno acumulado. Para motivos de comparação, as técnicas MV e HRP foram também analisadas sem o uso da pré-seleção, em comparação com a estratégia ingênua $1/N$ e o benchmark EEM.

Os resultados empíricos usando 10 ETFs internacionais que reproduzem os índices do mercado acionário de cada mercado emergente analisado em um período próximo de 8 anos de observações diárias mostram que as técnicas de pré-seleção por aprendizagem de máquina melhoram o desempenho das técnicas em relação ao retorno e o Índice de Sharpe, sendo as técnicas que envolvem otimização de portfólio de MV as que melhor performaram, com destaque especial para o KNN + MV. Em relação à volatilidade, a pré-seleção de ativos por *machine learning* não resultou em vantagens significativas em nenhuma das medidas estudadas.

O presente trabalho avançou no estudo empírico de modelos de portfólio de risco-retorno, integrando-os implicitamente com a previsão de retorno por meio da pré-seleção de ativos resultante dos algoritmos de *machine learning*. Além disso, visou-se dar maior ciência aos movimentos do mercado de capital dos países emergentes, ainda carente de estudos aprofundados. Pesquisas futuras podem comparar a eficácia da incorporação de vários mecanismos de previsão de retorno em diferentes modelos de portfólio e classes de ativos.



REFERÊNCIAS

REFERÊNCIAS

REFERÊNCIAS

- Adya, M., & Collopy, F. (1998). How effective are neural networks at forecasting and prediction? A review and evaluation. *Journal of Forecasting*, 17(1), pp. 481-495.
- Baek, Y., & Kim, H. Y. (2018). ModAugNet: A new forecasting framework for stock market index value with an overfitting prevention LSTM module and a prediction LSTM module. *Expert Systems With Applications*, 113, pp. 457-480. doi:10.1016/j.eswa.2018.07.019
- Ballings, M., Van den Poel, D., Hespeels, N., & Gryp, R. (2015). Evaluating multiple classifiers for stock price direction prediction. *Expert Systems with Applications*, 42(20), pp. 7046-7056. doi:10.1016/j.eswa.2015.05.013
- Ban, G.-Y., El Karoui, N., & Lim, A. E. (2018). Machine Learning and Portfolio Optimization. *Management Science*, 64(3), pp. 983-1476. doi:10.1287/mnsc.2016.2644
- Beale, E. M. (1959). On quadratic programming. *Naval Research Logistics Quarterly*, 6(3), pp. 227-243.
- Bentley, J. L. (1975). Multidimensional binary search trees used for associative searching. *Communications of the ACM*, 18(9), pp. 509-517.
- Bessler, W., Opfer, H., & Wolff, D. (2017). Multi-asset portfolio optimization and out-of-sample performance: an evaluation of Black-Litterman, mean-variance, and naïve diversification approaches. *The European Journal of Finance*, 23(1), pp. 1-30. doi:10.1080/1351847X.2014.953699
- Bishop, C. (1995). *Neural Networks for Pattern Recognition*. Oxford University Press.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45, pp. 5-32.
- Burggraf, T. (2021). Beyond risk parity – A machine learning-based hierarchical risk parity approach on cryptocurrencies. *Finance Research Letters*, 38. doi:10.1016/j.frl.2020.101523
- Chen, Y., & Hao, Y. (2018). Integrating principle component analysis and weighted support vector machine for stock trading signals prediction. *Neurocomputing*, 321, pp. 381-402. doi:10.1016/j.neucom.2018.08.077

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), pp. 273–297.

De Prado, M. L. (2016). Building Diversified Portfolios that Outperform Out of Sample. *The Journal of Portfolio Management*, pp. 59-69. doi:10.3905/jpm.2016.42.4.059

Dudoit, S., Fridlyand, J., & Speed, T. P. (2002). Comparison of discrimination methods for the classification of tumors using gene expression data. *Journal of the American Statistical Association*, 97, pp. 77-87.

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions networks for financial market predictions. *European Journal of Operational Research*, 270(2), pp. 654-669. doi:10.1016/j.ejor.2017.11.054

Frank, M., & Wolfe, P. (1956). An algorithm for quadratic programming. *Naval Research Logistics Quarterly*, 3, pp. 95–110.

Goetzmann, N. W., Li, L., & Rouwenhorst, K. G. (Janeiro de 2005). Long-Term Global Market Correlations. *The Journal of Business*, 78(1), pp. 1-38. doi:10.1086/426518

Grossman, S. J., & Zhou, Z. (1993). Optimal Investment Strategies for Controlling Drawdowns. *Mathematical Finance*, 3(3), pp. 241-276.

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies*, 33, pp. 2223–2273. doi:10.1093/rfs/hhaa009

Guerard, J. B., Markowitz, H., & Xu, G. (2015). Earnings forecasting in a global stock selection model and efficient portfolio construction and management. *International Journal of Forecasting*, 31(2), pp. 550-560. doi:10.1016/j.ijforecast.2014.10.003

Ho, T. K. (1995). Random decision forests. Proceedings of the 3rd International Conference on Document Analysis and Recognition, pp. 278-282.

Huang, C.-F. (2012). A hybrid stock selection model using genetic algorithms and support vector regression. *Applied Soft Computing*, 12(2), pp. 807-818. doi:10.1016/j.asoc.2011.10.009

Kara, Y., Boyacioglu, M. A., & Baykan, Ö. K. (2011). Predicting direction of stock price index movement using artificial neural networks and

support vector machines: The sample of the Istanbul Stock Exchange. *Expert Systems with Applications*, 5, pp. 5311-5319.

Långkvist, M., Karlsson, L., & Loutfi, A. (2014). A review of unsupervised feature learning and deep learning. *Pattern Recognition Letters*, 42, pp. 11-24. doi:10.1016/j.patrec.2014.01.008

LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., & Hubbard, W. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1(4), pp. 541-551.

Ma, Y., Han, R., & Wang, W. (2021). Portfolio optimization with return prediction using deep learning and machine learning. *Expert Systems With Applications*, 165, pp. 1-15. doi:10.1016/j.eswa.2020.113973

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, 7(1), pp. 77-91.

Markowitz, H. (1956). The optimization of a quadratic function subject to linear constraints. *Naval Research Logistics*, 3(1-2), pp. 111-133. doi:10.1002/nav.3800030110

Michaud, R. O. (1989). The Markowitz Optimization Enigma: Is 'Optimized' Optimal? *Financial Analysts Journal*, pp. 31-42. doi:10.2139/ssrn.2387669

Paiva, F. D., Cardoso, R. T., Hanaoka, G. P., & Duarte, W. M. (2019). Decision-making for financial trading: A fusion approach of machine learning and portfolio selection. *Expert Systems with Applications*, 115, pp. 635-655. doi:10.1016/j.eswa.2018.08.003

Perrin, S., & Roncalli, T. (2020). Machine Learning Optimization Algorithms & Portfolio Allocation. Em E. Jurczenko (Ed.), *Machine Learning for Asset Management: New Developments and Financial Applications* (Vol. 1, pp. 261-328). ISTE Ltd and John Wiley & Sons, Inc. doi:10.1002/9781119751182.ch8

Raffinot, T. (2018). Hierarchical Clustering-Based Asset Allocation. *The Journal of Portfolio Management*, 2, pp. 89-99. doi:10.3905/jpm.2018.44.2.089

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(3), pp. 386-408. doi:10.1037/h0042519

Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society*, 58(1), pp. 267–288.

Wang, W., Li, W., Zhang, N., & Liu, K. (2020). Portfolio formation with preselection using deep learning from long-term financial data. *Expert Systems With Applications*, 143, pp. 1-17. doi:10.1016/j.eswa.2019.113042

Wolfe, P. (1959). The simplex method for quadratic programming. *Econometrica*, 27(3), pp. 382-398.

Xiao, Y., Xiao, J., Lu, F., & Wang, S. (2013). Ensemble ANNs-PSO-GA approach for day-ahead stock e-exchange prices forecasting. *International Journal of Computational Intelligence Systems*, 6(1), pp. 96-114.

Yu, J. R., Chiou, W. J., Lee, W. Y., & Lin, S. J. (2020). Portfolio models with return forecasting and transaction costs. *International Review of Economics and Finance*, 66, pp. 118-130. doi:10.1016/j.iref.2019.11.002



idn

Bo
pro
cit
ref
Nos
são

idp

A ESCOLHA QUE
TRANSFORMA
O SEU CONHECIMENTO