

CARTEIRAS DE BLACK LITTERMAN COM ANÁLISES BASEADAS EM REDES NEURAIAS

Diego Guerreiro Bernardes

Depto. Eng. de Telecomunicações e Controle/EPUSP
diegogbernardes@usp.br

Oswaldo Luiz do Valle Costa 

Depto. Eng. de Telecomunicações e Controle/EPUSP
oswaldo@lac.usp.br

Resumo – Este artigo apresenta um sistema autônomo de gestão de carteiras. Sistemas autônomos de investimento consistem de regras para compra e venda de ativos no mercado financeiro que podem ser executadas por máquinas, com o objetivo de maximizar o retorno do investidor. O sistema utiliza Redes Neurais Artificiais para monitoramento do mercado e o modelo de Black-Litterman para otimização da alocação de patrimônio. O sistema analisa as dez ações mais negociadas do índice Bovespa, com redes neurais dedicadas a cada ação, e prevê estimativas de variações de preços para um dia no futuro a partir de indicadores da análise técnica. As estimativas das redes são então inseridas em um otimizador de carteiras, que utiliza o modelo de Black-Litterman, baseado em inferência Bayesiana, para compor carteiras diárias que empregam a estratégia *Long and Short*. Os resultados obtidos são comparados a um segundo sistema de trading autônomo, denominado Benchmark, sem o emprego da otimização de carteiras. Foram observados resultados com ótimo índice de Sharpe do modelo proposto em comparação ao Benchmark. Tais resultados sugerem que a utilização de modelos de inferência bayesiana combinados com redes neurais podem ser uma boa alternativa na gestão de carteiras.

Palavras-chave – Otimização de Carteiras, Modelo de Black Litterman, Inferência Bayesiana, Redes Neurais Artificiais, Perceptron Multi Camadas, Hedge.

Abstract – This paper presents an autonomous portfolio management system. Autonomous investment systems consists of a serie of buy and sell rules on financial markets, which can be executed by machines, oriented to maximizing investor gains. The system uses a Neural Network approach for monitoring the market and the Black-Litterman model for portfolio composition. The ten most traded assets from the Bovespa Index are analyzed, with dedicated neural networks, which suggests future return estimates using technical indicators as input. Those estimates are inserted in the Black-Litterman model which proposes daily portfolio composition using long & short positions. The results are compared to a second autonomous trading system without the Black-Litterman approach, referred to as Benchmark. The numerical results show a great performance compared to the Benchmark, specially the risk-return ratio, captured by the Sharpe Index. Such results suggest that the use of Bayesian inference models combined with neural networks may be a good alternative in portfolio management.

Keywords – Portfolio Optimization, Black-Litterman Model, Bayesian Inference, Artificial Neural Networks, MultiLayer Perceptron, Hedging strategies.

1 Introdução

O modelo de Black Litterman, estabelecido em Black and Litterman, 1991 e Black and Litterman, 1992 busca reduzir a distância existente entre as aplicações práticas de otimização por média e variância e sua proposta acadêmica, inicialmente feita por Markowitz, 1952. Otimizações de carteira, em geral, buscam aumentar o retorno do investidor para cada unidade de risco assumida, ou reduzir o risco para um determinado retorno esperado fixo. A técnica de otimização por média e variância considera o desvio padrão no histórico de preços de um ativo como unidade de risco, e busca construir carteiras de investimento equilibrando riscos assumidos ao fazer a combinação de ativos com covariância negativa, priorizando ativos com retorno esperado positivo. Nos artigos de He and Litterman, 2002 e Michaud, 1989, os autores argumentam sobre as limitações da otimização sem restrições que, por vezes, produz carteiras impraticáveis e contra-intuitivas, com pouca diversidade, alta alavancagem e alta sensibilidade aos erros nas estimativas de retorno esperado. A boa especificação de restrições e estimativa dos retornos esperados, portanto, muitas vezes se mostra uma atividade que demanda bastante esforço na aplicação prática do modelo de Markowitz. A grande contribuição do modelo de Black-Litterman está na capacidade de combinar as interpretações de gestores de investimentos sobre o mercado com os retornos esperados do modelo de Markowitz, criando um mecanismo mais intuitivo para o cálculo das estimativas dos retornos esperados dos ativos.

Esta abordagem é muito condizente com os objetivos buscados pelos sistemas de *soft-computing* aplicados às finanças, como as técnicas de aprendizado de máquina aplicadas à predição de variações de preços de ativos financeiros, especialmente as redes

neurais artificiais.

Um sistema de *trading* consiste num conjunto de regras para compra e venda de ativos com o objetivo de maximizar o rendimento do investidor. A proposta do sistema de *trading* apresentado neste artigo é uma abordagem quantitativa do modelo de Black-Litterman utilizando *soft-computing*. Este sistema monitora o comportamento diário das dez ações de maior peso do índice bovespa (Ibovespa) e infere, a partir de um conjunto de redes neurais artificiais dedicadas, alimentadas por entradas baseadas em indicadores da análise técnica utilizados por investidores de estratégias seguidoras de tendência, previsões de retornos esperados. Estas previsões são inseridas em um núcleo construído a partir do modelo de Black Litterman, que sugere uma carteira casada de compra e venda (*Long & Short*). Uma carteira (*Long & Short*) busca diminuir os riscos do investidor ao comprar ativos com expectativa de alta e vender ativos com expectativa de baixa. A performance do sistema é comparada com um segundo sistema de trading, também utilizando redes neurais construídas com a mesma metodologia, que negocia diretamente o índice Bovespa sem a utilização da otimização de carteiras. Assim, busca-se contribuir na apresentação de evidências a favor do impacto positivo da utilização de inferência Bayesiana, como o Modelo de Black-Litterman, junto à técnicas quantitativas de gestão de patrimônio. O artigo está organizado da seguinte forma. Na seção 2 é apresentada uma breve revisão bibliográfica deste trabalho. A seção 3 apresenta o modelo de Black-Litterman, as variações mais utilizadas na literatura e a base teórica utilizada no modelo. Na seção 4 é descrito o funcionamento do sistema, alguns resultados parciais e o detalhamento da construção do sistema. Finalmente, na seção 5, são apresentadas as métricas de desempenho do sistema e os resultados comparativos com o *Benchmark* seguida de uma breve discussão sobre os resultados obtidos e sugestões de aprimoramento do sistema na seção 6.

2 Revisão Bibliográfica

Nos trabalhos de Black and Litterman, 1991 e Black and Litterman, 1992 são estabelecidas as bases conceituais do que posteriormente viria a ser conhecido como o modelo de Black-Litterman. O modelo é uma proposta complementar às técnicas de otimização por média e variância que, segundo os autores, apresentam dificuldades de implementação que as distanciam da realidade do mercado.

A proposta dos autores utiliza o conceito de “visões do investidor” que, somadas aos retornos esperados de equilíbrio de uma carteira de referência, produz uma nova estimativa do retorno esperado, que leva a uma nova carteira otimizada a partir dos conceitos de covariância empregados na técnica de Markowitz, 1952. As visões do investidor compõem um mecanismo por onde o administrador de um portfólio pode manifestar sua interpretação sobre as tendências do mercado, ou combinar as interpretações de analistas, utilizando um conjunto de matrizes. Ao se aplicar a otimização reversa, técnica derivada a partir da otimização de carteiras de Sharpe, 1964, obtém-se uma aproximação dos retornos utilizados naquela otimização, que podem ser considerados retornos de equilíbrio.

O mecanismo de combinação entre a carteira de equilíbrio e a carteira de visões do investidor utiliza um parâmetro de confiança para cada expressão de visão. Ou seja, na ausência de visões, ou caso o investidor não possua uma opinião formada sobre um ou mais ativos, é utilizado o retorno de equilíbrio para a composição da carteira, que se torna a própria carteira de equilíbrio. O desvio entre o retorno de equilíbrio e o retorno *a posteriori*, chamado de retorno de Black-Litterman, relaciona-se com o nível de confiança, ou erro, que o investidor apresenta para sua estimativa. Entretanto, as publicações originais do modelo não apresentam todo o desenvolvimento matemático para sua implantação, apresentados em trabalhos posteriores, fato este que poderia explicar o baixo número de trabalhos relacionados ao modelo ao longo da década de 1990.

A influência do modelo de Black-Litterman aumentou nas décadas seguintes, em grande parte devido a outros trabalhos ao final da década de 1990 e começo da década de 2000, como He and Litterman, 2002, Satchell and Scowcroft, 2000, Lee, 2000 e Idzorek, 2007 que apresentam revisões e implementações do modelo original, bem como desenvolvimentos complementares, propondo alternativas de abordagem no cálculo de alguns parâmetros, e mostrando a intuição por trás dos conceitos utilizados na construção do modelo.

Idzorek, 2007, além de apresentar uma implementação passo a passo do modelo, propõe uma abordagem mais intuitiva para o cálculo da matriz de variância Ω , utilizando um índice de confiança entre 0% e 100% para cada visão do investidor.

Fabozzi et al., 2006 implementa uma estratégia de *trading* utilizando o modelo de Black-Litterman. É reforçado que muitas das estratégias de *trading* utilizadas hoje dificilmente podem ser traduzidas em previsões de retornos esperados, como requer a otimização por média e variância clássica ou o modelo de Black-Litterman. A estratégia sugerida pelo autor aloca pesos positivos (compra) para os ativos que pertencem à metade superior de um ranking, composto por ativos com melhor retorno nos últimos períodos, e pesos negativos (venda) aos ativos de pior retorno.

Diversos trabalhos nas últimas década exploraram as potencialidades das redes neurais em sistemas preditivos no mercado financeiro. Trabalhos pioneiros, como Kimoto et al., 1990, White, 1988, Kaastra and Boyd, 1996, criaram as bases para sistemas mais recentes de predição utilizando ferramentas da análise técnica ou sistemas de trading, como Gambogi, 2013 e Klassen, 2007.

Embora muitos trabalhos utilizando o modelo de Black-Litterman em gestão ativa de carteiras foram encontrados, poucos destes trabalhos utilizam sistemas de redes neurais artificiais em conjunto com esta abordagem, proposta aqui apresentada. De maneira bastante interessante, o trabalho de Xing et al., 2018 apresenta um sistema composto de redes neurais para monitoramento de opinião pública através de redes sociais na composição de visões de mercado utilizadas no modelo de Black-Litterman. Em Wutsqa and Kusumawati, 2016 e Zimmermann et al., 2001 são apresentados modelos de previsão de preços para composição das visões de Black-Litterman, assim como neste trabalho, porém sem a utilização de indicadores de análise técnica. A maior contribuição apresentada neste artigo, em relação aos outros trabalhos publicados, é a utilização de um sistema de trading diário

que combina redes neurais com a formulação de Black-Litterman e que pode ser implementado para qualquer ativo financeiro e indicador de análise técnica como entrada do modelo.

3 Modelo de Black Litterman

O modelo de Black-Litterman estima o retorno esperado de ativos financeiros utilizando inferência Bayesiana para combinar o equilíbrio do mercado no longo prazo com as visões dos investidores sobre os retornos futuros. Para possibilitar a construção do modelo, assume-se que os retornos dos ativos comportam-se como variáveis aleatórias com distribuições normais de média desconhecida e variância conhecida. Algumas liberdades de implementação do modelo podem ser encontradas na literatura, sendo duas instâncias as mais recorrentes. A implementação original, de Black and Litterman, 1992 e He and Litterman, 2002 por vezes chamada de modelo de referência, e uma implementação utilizando abordagem de encolhimento (*Shrinkage*) observada nos modelos de Satchell and Scowcroft, 2000 e Meucci, 2006, aqui chamada de modelo alternativo.

A teoria moderna de carteiras, consagrada no trabalho de Markowitz, 1952, considera o vetor de retorno (r) de um conjunto de ativos como um vetor aleatório com distribuição de probabilidade gaussiana com valor esperado (μ) e matriz de variância (Σ), ou de forma mais compacta $r \sim N(\mu, \Sigma)$. O modelo de referência considera o valor esperado (μ) da distribuição r também uma variável aleatória, de distribuição gaussiana desconhecida passível de inferência. Na inferência bayesiana parte-se de uma premissa inicial, conhecida *a priori*, e infere-se um novo valor *a posteriori* a partir de informações adquiridas ou evidências encontradas. Para a estimação de μ parte-se do valor de equilíbrio calculado pelo método CAPM (*Capital Asset Pricing Model*), definido em Sharpe, 1964, e agrega-se as informações das crenças – subjetivas ou quantitativas – dos agentes do mercado, chamadas de visões dos investidores no trabalho original, como novas informações. O modelo de Black-Litterman assume, portanto, que o preço de um ativo é definido por duas fontes de informações – o prêmio de risco no equilíbrio de mercado, e as visões dos investidores.

No equilíbrio de mercado, uma abstração matemática, todas as demandas e ofertas por um ativo são igualadas e os investidores que mantêm um ativo esperam um retorno definido pelo prêmio de risco de equilíbrio (Π). Uma maneira de estimar o prêmio de risco é assumir que todos os agentes de mercado utilizam a metodologia do CAPM. Assim, pode-se calcular o prêmio de risco através de uma abordagem inversa à do CAPM.

O cálculo do valor de μ parte deste conceito. O valor *a priori* de μ – uma variável aleatória não observável – seria, portanto, o valor do prêmio de risco de equilíbrio mais um valor de erro de estimativa oriundo da incerteza sobre o CAPM e o equilíbrio. Ou seja:

$$\mu = \Pi + \epsilon^e$$

Com ϵ^e sendo um vetor aleatório de média igual a zero e matriz de covariância igual a $\tau\Sigma$. Pode-se notar que τ , neste caso, é um escalar que denota a incerteza em cima da estimativa *a priori*, ou a confiança no CAPM – incerteza esta de alguma maneira proporcional à sua covariância Σ . É justamente esta variável τ que protagoniza a principal diferença entre os modelos de referência e alternativo. No modelo de referência o valor esperado do retorno de um ativo segue uma distribuição gaussiana dada por $\mu \sim N(\Pi, \tau\Sigma)$.

O modelo alternativo considera o valor esperado μ não como uma variável aleatória com distribuição gaussiana, mas como uma estimativa pontual. Implica-se, então, que a informação *a priori* da inferência Bayesiana iguala-se ao prêmio de risco de equilíbrio, ϵ^e iguala-se a zero e a variável τ , ainda presente, adquire um significado de variável de controle no modelo. Neste novo contexto a variável τ ajusta quanto as visões do investidor devem influenciar no valor *a posteriori* do retorno do ativo, ou seja, ela pode ser considerada como uma variável de confiabilidade nas visões dos investidores. Como observado por Walters, 2014, existem algumas implicações nesta interpretação de τ , discutidas adiante.

3.1 O Modelo de Referência

Na equação (1), equação de referência do modelo de Black-Litterman, os retornos estimados *a posteriori* são representados pela matriz de retornos $\hat{\Pi}$, de dimensão $N \times 1$, e pela matriz de covariância M obtida pela equação (2), de dimensões $N \times N$.

$$\hat{\Pi} = [(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1}[(\tau\Sigma)^{-1}\Pi + P'\Omega^{-1}Q] \quad (1)$$

$$M = \Sigma + [(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1} \quad (2)$$

Nestas equações as informações *a priori* são expressas na forma da matriz de covariância $\tau\Sigma$ e no prêmio de risco de equilíbrio Π .

O prêmio de risco Π pode ser obtido pela otimização reversa da função utilidade quadrática do investidor – que resulta em $\Pi = \delta\Sigma w$, com w sendo o conjunto de pesos de uma carteira e δ o coeficiente de aversão ao risco desta carteira.

As informações parciais sobre o mercado são representadas pelas matrizes P , Q e Ω , que compõem as visões dos investidores. Estas visões são impressões, crenças ou constatações quantitativas do investidor sobre o mercado e são expressas como retornos esperados. Estes retornos podem ser na forma absoluta, onde o investidor estima a oscilação futura de um ativo, na forma relativa, onde o investidor estima a oscilação de um ativo em relação a outro(s), ou uma combinação destas situações.

Seja um mercado de n ativos, onde expressam-se k visões, as matrizes das visões do investidor podem ser entendidas na forma de uma carteira, cujo retorno pode ser descrito pela equação (3).

$$P\mu = Q + \epsilon^\nu \quad (3)$$

Sendo μ , um vetor com dimensões $n \times 1$, o retorno esperado da carteira que as visões exprimem, P a matriz de pesos, dimensões $k \times n$, que compõe a carteira de visões, Q a matriz de estimativas de retornos do investidor, de dimensões $k \times 1$, e ϵ^ν indica a incerteza nas estimativas de Q . Caso o investidor expresse uma visão na forma absoluta, a soma dos pesos da linha da matriz P para esta visão deve ser igual à 1, e para a forma relativa a soma dos pesos deve ser igual à 0.

O componente ϵ^ν , neste caso, pode ser interpretado como a incerteza na especificação da visão. Este pode ser representado por um vetor de ruído branco de média zero e variância finita e conhecida, igual à Ω , de dimensões $k \times k$. Embora não seja uma exigência matemática, costuma-se expressar os componentes da matriz Ω (ω_i) como independentes e não correlacionadas, tornando-a uma matriz diagonal. Esta simplificação ajuda na especificação de Ω pois, caso contrário, o investidor teria que estimar não só a incerteza na sua expectativa como também a incerteza na correlação entre as suas expectativas, tornando o problema bastante complexo.

Por exemplo:

Considere um mercado de quatro ativos onde o investidor deseja manifestar duas opiniões. A primeira opinião, manifestada de forma relativa, diz que o ativo 1 terá uma performance 2% melhor que o ativo 3, com uma confiabilidade ω_1 . A segunda opinião, expressa em forma absoluta, diz que o ativo 2 irá desvalorizar 6%, com uma confiabilidade ω_2 . Desta forma, temos as três matrizes a seguir:

$$P = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}; Q = \begin{bmatrix} 3 \\ -6 \end{bmatrix}; \Omega = \begin{bmatrix} \omega_1 & 0 \\ 0 & \omega_2 \end{bmatrix}$$

3.2 O modelo alternativo

A equação (1) do retorno esperado *a posteriori* pode ser reescrita na equação (4), onde é possível observar como as visões dos investidores formam um segundo componente que é adicionado ao retorno esperado *a priori*.

$$\hat{\Pi} = \Pi + \tau \Sigma P' [(P \tau \Sigma P') + \Omega]^{-1} [Q - P \Pi] \quad (4)$$

Analisando a equação (4) nota-se que para $|\Omega| \rightarrow 0$, o retorno *a posteriori* se iguala aos da carteira de visão do investidor – equação (3) – e, para $|\Omega| \rightarrow \infty$, estes retornos se igualam ao prêmio de equilíbrio Π . Ou seja, Ω é uma grandeza inversamente proporcional à confiança do investidor em suas estimativas, e pode ser interpretado como a covariância das visões do investidor. A abordagem para o cálculo de Ω é aberta à discussão. Em He and Litterman, 2002 considera-se que esta covariância é proporcional à covariância *a priori*, como segue:

$$\begin{aligned} \omega_{ij} &= p_i (\tau \Sigma) p_j^T, \quad \forall i = j \\ \omega_{ij} &= 0, \quad \forall i \neq j \\ &\text{ou} \\ \Omega &= \text{diag}(P(\tau \Sigma)P') \end{aligned} \quad (5)$$

Onde p_i são componentes da matriz P e ω_{ij} os componentes da matriz Ω . O modelo alternativo, como descrito anteriormente, difere do modelo de referência pela simplificação de μ para uma estimação pontual, igualando-o ao prêmio de risco de equilíbrio Π . A consequência primária desta simplificação é a perda de significado da equação (2) – pois $\hat{\Pi}$ torna-se também uma estimação pontual – seguida da mudança no significado da variável τ . Enquanto no modelo de referência τ é relacionada à incerteza do CAPM e do resultado da otimização reversa, no modelo alternativo τ torna-se uma variável de controle da combinação entre o prêmio de equilíbrio e as visões do investidor, e sua função confunde-se com a da matriz Ω . Dado este conflito, no modelo alternativo pode-se igualar τ à 1 e controlar a combinação de retornos apenas através de Ω .

Utilizando o modelo alternativo, Idzorek, 2007 propõe um método intuitivo para a estimativa de Ω , onde o investidor estabelece um nível de confiança entre 0% e 100% para cada visão.

Em Walters, 2014, o autor observa que este método pode ser simplificado adicionando uma variável de encolhimento α à equação (5). Assim, nesta abordagem com $\tau = 1$, Ω pode ser interpretado como uma variável de encolhimento, que passa a ser controlado por α através da equação reescrita (6), com c , um escalar, sendo o nível de confiança entre 0% e 100%.

$$\begin{aligned} \Omega &= \alpha \text{diag}(P \Sigma P') \\ \alpha &= \frac{1 - c}{c} \end{aligned} \quad (6)$$

4 Modelo Proposto e Metodologia

O objetivo do modelo proposto é buscar evidências favoráveis à utilização da abordagem bayesiana em sistemas autônomos de investimento e gestão de patrimônio. Para este fim foi construído um sistema composto por dois núcleos. O primeiro núcleo é responsável pelo monitoramento do mercado utilizando redes neurais. O segundo núcleo trata da inferência bayesiana em si, através do modelo de Black-Litterman, utilizando as análises do primeiro núcleo para sugerir uma carteira de posições casadas de compra e venda para as ações monitoradas. Uma operação de compra consistem em adquirir ativos, e assim ter uma exposição positiva às suas variações de preços, uma opção de venda consiste em tomar emprestado um ativo de um terceiro e vendê-lo no mercado com o compromisso de reverter a operação no futuro, ficando exposto negativamente às suas variações de preços. Os resultados do modelo são então comparados aos de um segundo sistema autônomo de investimento, sem a utilização da inferência bayesiana, e as métricas de desempenho são analisadas sob uma ótica de risco e retorno.

O modelo foi implementado utilizando ferramentas *open-source*. Utilizou-se a API Keras, Chollet et al., 2015, com a *engine* TensorFlow para a construção das redes neurais e as bibliotecas SciPy, Jones et al., 2001 e NumPy, Oliphant, 2006, para manipulação de dados.

Com o objetivo de diminuir as exigências computacionais, o sistema proposto monitora um mercado com número reduzido de ativos. Para tanto, foram escolhidas as dez ações com maior peso do índice Bovespa ao longo do período analisado, de 30/04/2014 à 29/04/2016. Os ativos escolhidos são: Ambev SA [ABEV3], Banco Bradesco SA [BBDC4], BRF S.A. [BRFS3], Cielo SA [CIEL3], Itaúsa Investimentos Itaú SA [ITSA4], Petrobrás ON [PETR3], Petrobrás PN [PETR4], Ultrapar Participações SA [UGPA3], e Vale [VALE3].

O sistema funciona sob as seguintes regras: As posições são sugerida ao final do dia, após o fechamento do mercado, para o pregão do dia seguinte. Utilizam-se como entradas a carteira Ibovespa – que é conjunto de ativos e suas devidas proporções na composição do índice – do final do pregão para o cálculo dos retornos de equilíbrio, e as previsões de preço de fechamento para o d_{i+1} , feitas pelo conjunto de redes neurais do primeiro núcleo transformadas em previsões de retornos. No início do pregão seguinte montam-se as operações e ao final do pregão vende-se todos os ativos. Então, novamente é feita a previsão de preços do dia seguinte, e as etapas se repetem. A figura 1 ilustra o funcionamento diário do sistema de trading de maneira simplificada.

Obviamente, uma atualização diária da carteira resulta em elevados custos operacionais junto às corretoras de ações em um ambiente real. Como o objetivo deste sistema é analisar o modelo de Black-Litterman utilizado em conjunto com as previsões de um sistema quantitativo autônomo, os custos operacionais foram desconsiderados.

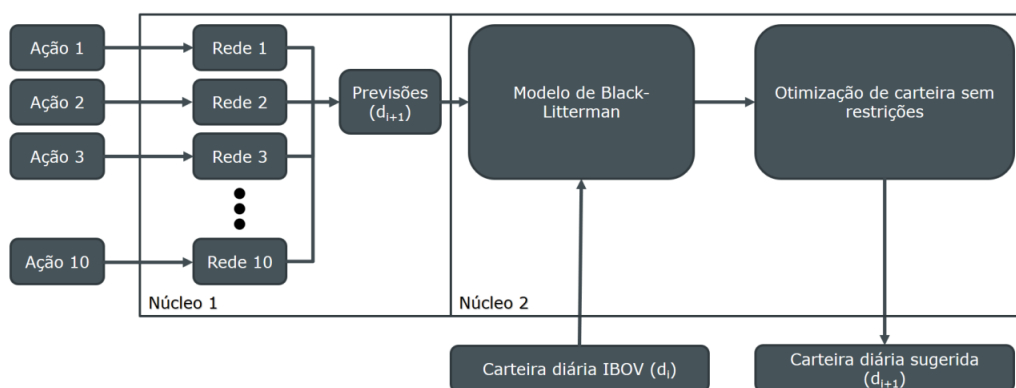


Figura 1: Diagrama Simplificado do Sistema

4.1 Primeiro Núcleo: Implantação das Redes Neurais

A função do primeiro núcleo do sistema é acompanhar e monitorar os ativos no mercado, e oferecer um conjunto de previsões de oscilações para o segundo núcleo. O primeiro núcleo é composto por um conjunto de dez redes neurais. As redes neurais foram construídas baseadas em resultados de trabalhos anteriores, e suas arquiteturas foram definidas por um processo seletivo utilizando busca exaustiva. Neste processo seletivo, treinou-se o maior número possível de redes neurais por ativo, utilizando-se um conjunto de dados, e foram selecionadas as arquiteturas de redes que frequentemente produzissem boa performance, quando utilizadas para prever oscilações em outro conjunto de dados. Um terceiro conjunto de dados – onde o modelo proposto foi efetivamente testado – foi apresentado apenas para as arquiteturas vencedoras, e apenas após a etapa de retreinamento utilizando janelamento móvel. A base de dados consiste em históricos de preços diários das ações analisadas, trabalhados utilizando indicadores de análise técnica. A figura 2 ilustra as etapas para implantação de uma rede neural do primeiro núcleo.

Buscou-se implementar modelos de redes neurais com bons resultados em trabalhos anteriores. Seguindo orientações de Kaastra and Boyd, 1996, Klassen, 2007 e Gambogi and Costa, 2013, as redes implementadas são variações do *Feed-Forward Multi-Layer Perceptron*, com uma camada escondida. Todas as redes foram treinadas utilizando treinamento supervisionado com a técnica de *cross-validation* para diminuir o problema de *overfitting*. Utilizou-se a função de transferência tangente hiperbólica

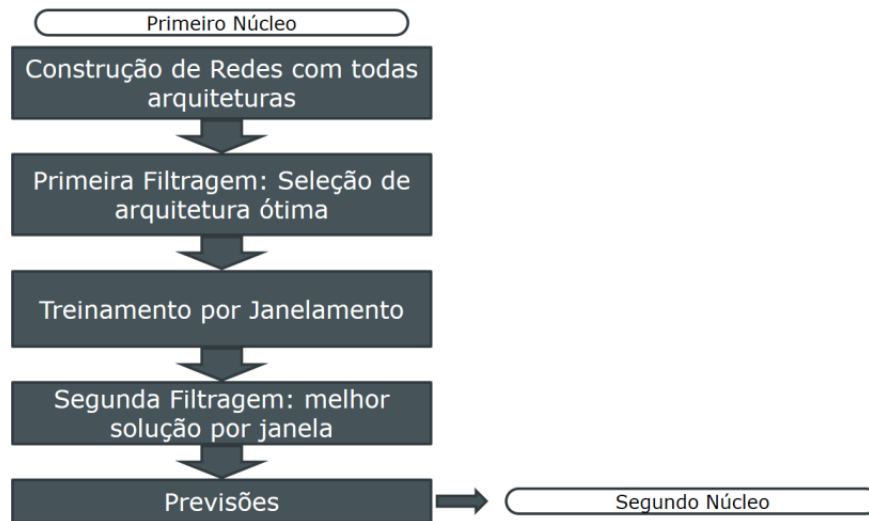


Figura 2: Etapas de implantação do primeiro núcleo

nas camadas escondidas, com funções lineares na camada de saída. O universo de possíveis entradas e arquiteturas foi reduzido a partir de busca exaustiva, até restar uma arquitetura vencedora por ativo, como descrito adiante.

4.1.1 Indicadores da Análise Técnica

Para compor o universo de entradas, foram escolhidos os indicadores de análise técnica utilizados em Klassen, 2007, com as modificações sugeridas por Gambogi and Costa, 2013. São eles: MACD (*Mean Average Convergente Divergence*) com períodos de 8 e 17 dias, ROC (*Rate of Change*) com período de 10 dias, além dos preços de fechamento do dia e dos últimos três dias ($d_i, d_{i-1}, d_{i-2}, d_{i-3}$). As redes devem fazer a previsão de um dia no futuro, portanto cada rede possui uma única saída, correspondente à previsão de preço em (d_{i+1}).

O indicador MACD, descrito em Murphy, 1999, busca captar modificações de momentum e direção de variações de preços. Ele é obtido através da diferença de uma média móvel 'rápida' e uma média móvel 'lenta'. A interpretação entre os analistas técnicos desta diferença é que os preços tendem a se manter próximos à média lenta, e consequentemente esta diferença constitui uma oportunidade de arbitragem. Embora esta interpretação seja irrelevante para a rede neural, ao aplicá-la espera-se que de algum maneira este comportamento esteja refletido em preços futuros. A equação (7) representa o cálculo do indicador. O indicador ROC, equação (8), também descrito em Murphy, 1999 e acrescentado em Gambogi and Costa, 2013, é um indicador puro de momentum e registra a variação de preço ao longo de um período.

$$MACD = MME(p_1) - MME(p_2) \quad (7)$$

$$ROC(p) = \frac{d_i - d_{i-p}}{d_{i-p}} \quad (8)$$

onde $MME(p)$ é a Média Móvel com decaimento Exponencial de período p , que pode ser implementada por:

$$MME(p) = \frac{P_p + (1 + \alpha)P_{p-1} + (1 + \alpha)^2P_{p-2} + \dots + (1 + \alpha)^pP_0}{1 + (1 + \alpha) + (1 + \alpha)^2 + \dots + (1 + \alpha)^p}$$

$$\alpha = \frac{2}{p + 1}$$

Com P_i sendo o preço de fechamento no dia i .

4.1.2 Conjuntos de Treinamento, Teste e Verificação

O conjunto total de dados utilizado foi dividido em três subconjuntos mutuamente exclusivos: treinamento, teste e verificação. A amostra de treinamento utiliza também a abordagem de treinamento/validação, que é outra sugestão para diminuição de *overfitting* e é nativa de algumas APIs, como o Keras. O modelo busca minimizar o erro na amostra de treinamento enquanto valida esta minimização em paralelo em uma amostra de validação. O processo de otimização é interrompido no ponto onde ocorre divergência entre os gradientes de erro de treinamento e validação. O primeiro e segundo conjunto, treinamento/validação e teste,

pertencem ao período de 30/04/2011 à 29/04/2014. Estes dados foram embaralhados, sendo que durante o treinamento as redes 'não sabiam' se tratar de uma série contínua. Após o embaralhamento, dividiu-se os dados efetivamente em treinamento/validação e teste em uma proporção de 80% e 20%. Após a finalização do treinamento as redes foram expostas ao conjunto de dados de teste para prever os preços em d_{i+1} . Utilizando estas previsões, foram selecionadas as arquiteturas que frequentemente produziram boas métricas de desempenho, segundo os critérios apresentados nas próximas seções.

O terceiro conjunto, verificação, correspondeu ao período de 30/04/2014 à 29/04/2016. O conjunto de verificação foi utilizado apenas com as redes vencedoras, e neste período o sistema foi efetivamente testado utilizando os dois núcleos.

As entradas das redes neurais foram normalizadas entre o intervalo de -0.75 à 0.75, como sugerido nos trabalhos relacionados, para estimular as redes a trabalhar dentro da região linear da função de transferência, minimizando os efeitos de valores muito grandes ou muito pequenos no cálculo do erro de *back-propagation*.

4.1.3 Seleção de Arquiteturas

A estratégia de maximizar as rodadas de treinamentos utilizando loops foi escolhida como opção para lidar com o problema dos mínimos locais utilizando busca exaustiva. O universo de arquiteturas possíveis foi reduzida utilizando abordagens de trabalhos anteriores. Assim, a filtragem partiu de dez arquiteturas definidas *a priori*, descritas na tabela 1.

	d_i	d_{i-1}	d_{i-2}	d_{i-3}	MACD	ROC	<i>hidden</i>
I	●	●	●	●	●	○	3
II	●	●	●	●	●	○	2
III	●	●	●	●	○	●	3
IV	●	●	●	●	○	●	2
V	●	●	●	○	●	○	2
VI	●	●	●	○	○	●	2
VII	●	●	○	○	●	○	1
VIII	●	●	○	○	○	●	1
IX	●	○	○	○	●	●	2
X	●	○	○	○	●	●	1

Tabela 1: Arquiteturas escolhidas *a priori*

Cada arquitetura possui um respectivo número de entradas, de camadas escondidas definidos pela variável *hidden* e todas possuem um neurônio na camada de saída. As camadas escondidas possuem função de transferência tangente hiperbólica e as camadas de saída possuem função de transferência linear. Para cada uma destas arquiteturas foram treinadas quinhentas redes. As redes foram aleatoriamente separadas em cinco séries de cem, e estas cinco séries foram paralelamente expostas à filtragem, sendo escolhida uma rede vitoriosa por série, como ilustra a figura 3. Ou seja, cinco redes vencedoras por arquitetura.

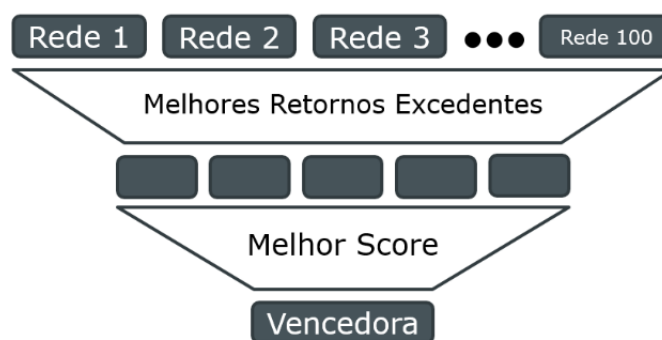


Figura 3: Filtragem por série de treinamentos

Como critério de filtragem levou-se em consideração três atributos de performance: O retorno excedente produzido pela rede, em comparação à estratégia de *buy & hold* no período de teste; A taxa de acertos na previsão de operações vencedoras no período (*score*), onde uma operação é contabilizada como vencedora quando a oscilação de preço do ativo ocorre na direção prevista pela rede, independente do preço previsto, e contabilizada como perdedora no caso contrário; e a raiz do erro médio quadrático (*RMSE: Root Mean Squared Error*) das previsões de preços. Uma estratégia *Buy & Hold* é a estratégia passiva padrão, onde o investidor compra um ativo e mantém sua posição comprada durante todo o período analisado.

A métrica de retornos excedentes produzido pela rede foi calculada como uma simulação da rede sendo utilizada automaticamente em uma estratégia de *trading*, utilizando-se a amostra de dados de teste. Nesta simulação, a rede prevê o preço futuro

em d_{i+1} e assume a posição comprada em caso de previsão de alta, ou assume posição vendida em caso de previsão de baixa. A métrica resultante é então o retorno excedente produzido, descontado do retorno passivo do mesmo ativo no período. Optou-se por utilizar o retorno excedente como indicador sobre o retorno absoluto da rede, para evitar falsos positivos em relação a ativos que tiveram um comportamento de ganho passivo no período. O *score* é calculado de maneira semelhante, contabilizando as operações vencedoras da rede. Embora o RMSE seja efetivamente a função que procurou-se minimizar durante o treinamento, a boa performance do sistema está relacionada à previsão correta de oscilações, motivo pelo qual o retorno excedente no período e o *score* da rede foram escolhidos como principais atributos de filtragem. O RMSE foi considerado como critério de desempate.

Após a aplicação do filtro, o universo de redes foi reduzido à cinco redes vencedoras por arquitetura. Calculou-se então a média das métricas (Retornos, Score, RMSE) destas redes e aplicou-se novamente a metodologia do filtro. A arquitetura vencedora por ativo, como ilustrado pela figura 4, foi aquela com melhor taxa de acerto médio entre as três arquiteturas de melhor retorno excedente médio, sendo o RMSE novamente utilizado como critério de desempate.

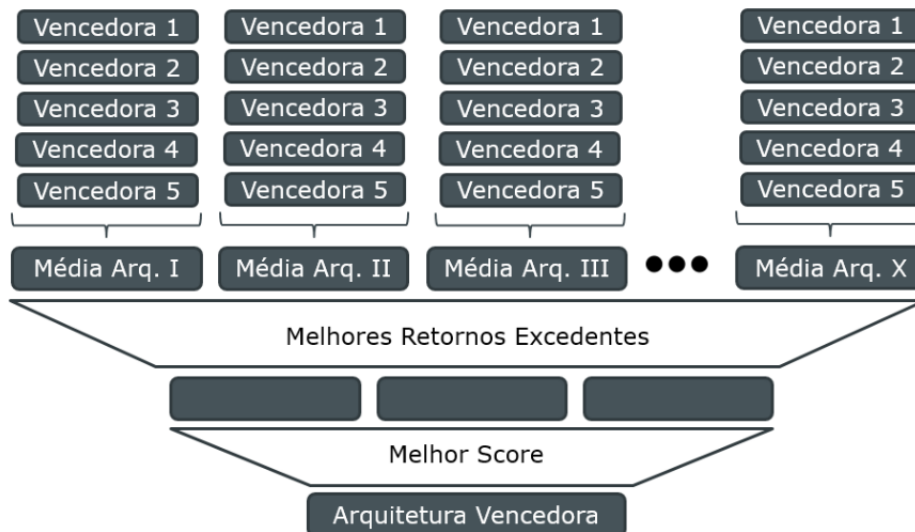


Figura 4: Filtragem de arquiteturas

Ativo	Arquitetura
ABEV3	VI
BBDC4	VII
BRFS3	X
CIEL3	VIII
ITSA4	V
ITUB4	X
PETR3	VII
PETR4	I
UGPA3	I
VALE3	III
IBOV	III

Tabela 2: Arquiteturas Vencedoras

O objetivo desta metodologia foi selecionar a arquitetura que resultasse em bons resultados com mais frequência, com os resultados descritos pela tabela 2, para posteriormente ser aperfeiçoada na etapa de janelamento móvel.

Ao longo deste processo, como pode ser observado na tabela 2, também foi treinada uma rede para o índice Bovespa [IBOV]. Esta rede foi posteriormente utilizada na construção do *Benchmark*, ilustrado mais adiante.

4.1.4 Janelamento Móvel

A abordagem de janelamento móvel busca retreinar a rede neural a determinados intervalos. Para isto, divide-se a amostra total de dados (treinamento, teste e verificação) pela quantidade de janelas desejadas, a cada início de período de verificação, utiliza-se o período imediatamente anterior para treinamento e testes, como ilustra a figura 5.

O embasamento desta abordagem encontra-se no caráter dinâmico e adaptativo do mercado financeiro, conhecido intuitivamente pelos operadores do mercado e presente em algumas formulações epistemológicas da formação de preços, como a hipótese

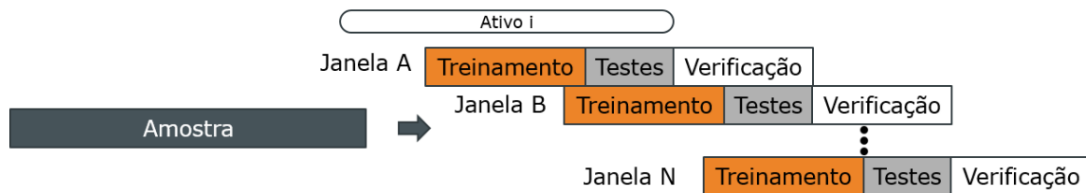


Figura 5: Janelamento

do mercado adaptativo, Lo, 2004 e a hipótese do mercado fractal Peters et al., 1994.

Note na tabela 2 que foi escolhida uma arquitetura para o índice IBOV. Embora o Ibovespa não seja efetivamente um ativo negociado em bolsa, sua variação de pontos foi tratada como um ativo financeiro nesta etapa, para o dimensionamento de um tamanho único de janelas comum a todos os ativos. A rede neural selecionada para compor o sistema de *Benchmark*, como será explicado adiante, também utilizou esta mesma arquitetura.

Os possíveis tamanhos de janelas foram definidos a partir do subconjunto de verificação. Este subconjunto foi composto por 494 amostras. Foram testadas janelas com dimensões $-(J)-$ de 13, 19, 26 e 38 amostras, múltiplos de 494. O tamanho das janelas de treinamento $-(JT)-$ deveriam ser números inteiros que maximizassem a utilização das amostras. O conjunto total de treinamento e teste era composto de 743 amostras, então a dimensão T foi definida pelo valor inteiro da divisão, como segue:

$$T = \left\lfloor \frac{743}{J} \right\rfloor$$

Para cada combinação de configuração de janela $J - T$ – utilizando-se apenas a amostra de treinamento e teste – aplicou-se novamente a metodologia de filtragem. Seguindo metodologia semelhante à etapa de seleção de arquitetura, foram treinadas 100 redes por janela, para cada configuração. Estas 100 redes foram separadas aleatoriamente em 5 séries de 20, das quais foram selecionadas 5 redes vencedoras por janela, utilizando as métricas de Retorno Excedente, Score e RMSE.

Janelas	Amostras de treinamento e teste	Amostras de verificação
38	741	13
26	741	19
19	728	26
13	722	38

Tabela 3: Configurações de Janelas

Assim, foram calculadas as médias destas 5 vencedoras por janela, resultando em um indicador de desempenho para cada janela. Naturalmente, a quantidade de janelas por configuração $J - T$ é variável, e conseqüentemente a quantidade de vencedoras por configuração é relacionada a esta grandeza, como mostra a tabela 3. Para poder comparar as diferentes configurações de janela, optou-se por tirar uma média destes indicadores, resultando na tabela 4. A configuração 38 – 19 obteve as melhores métricas médias.

J-T	Média de Retornos	Score Médio	RMSE Médio
13-57	8.71%	52.26%	1.23%
19-39	5.79%	51.47%	1.36%
26-28	6.31%	51.60%	1.33%
38-19	9.49%	52.33%	1.28%

Tabela 4: Métricas das Configurações de Janelas no período de Testes

4.1.5 Treinamento das Redes Neurais e Saída do Primeiro Núcleo

A última etapa do primeiro núcleo foi o treinamento das redes utilizando as dimensões de janelas definidas na etapa anterior. Como comentado anteriormente, isto significa um retreinamento contínuo a cada início de janela, como ilustra a figura 5.

Cada treinamento foi acompanhado de uma etapa de busca exaustiva de melhor conjunto de pesos. Para cada janela foram treinadas 100 redes neurais, separadas em 20 grupos de 5. Entre as 5 finalistas, foi escolhida uma rede vencedora. A figura 6 ilustra esta última etapa de filtragem. O produto desta etapa é um conjunto de pesos sinápticos por janela por ativo.

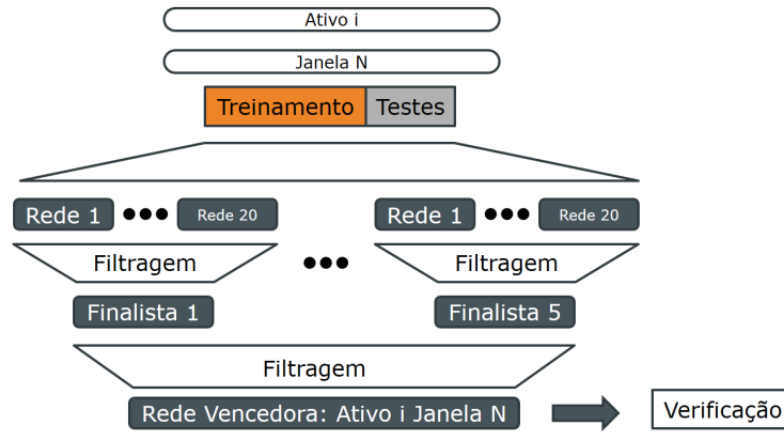


Figura 6: Treinamento de uma Janela

Como exemplo de resultado do primeiro núcleo, considere o ativo ABEV3. A arquitetura ótima encontrada foi a de número VI, repetida na tabela 5 para facilitar a visualização, e que também pode ser representada pelo diagrama da figura 7.

	d_i	d_{i-1}	d_{i-2}	d_{i-3}	MACD	ROC	hidden
VI	•	•	•	○	○	•	2

Tabela 5: Arquitetura ótima: ABEV3

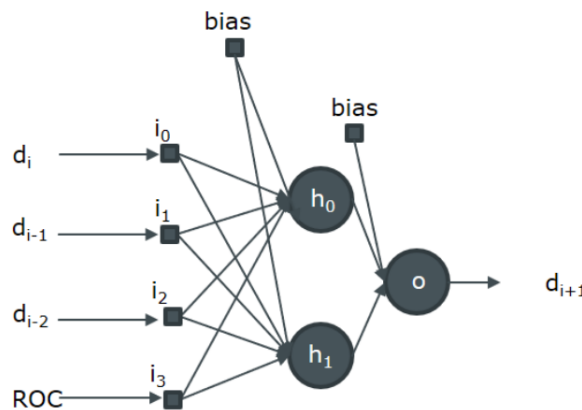


Figura 7: ABEV3

Os dados da primeira janela de treinamento/teste, limitada pelo período de 31/05/2011 a 29/04/2014, foram divididos de forma que 80% dos dados foram utilizados para treinamento e os outros 20% foram separados. Após treinamento foram obtidos 100 conjuntos de pesos sinápticos na forma da tabela 6.

	$[h_0, h_1]$		$[o]$
i_0	[0.66642487, -0.86955923]	h_0	[0.9292472]
i_1	[0.27950108, -0.695755]	h_1	[-1.25893271]
i_2	[0.88129503, -0.58044165]	bias	[0.13739333]
i_3	[-0.55097717, -0.62097532]		
bias	[0.09154823, -0.05482146]		

Tabela 6: Tabela de pesos sinápticos para ABEV3 - Janela 1

Então, utilizando os dados de treinamento – 20% da amostra – alimentou-se estas 100 redes em paralelo, e analisou-se a saída obtida \hat{y}_s em comparação com a saída conhecida y_s . Destas 100 redes, foi escolhida aquela com melhores métricas, cujos valores são estes da tabela 6. Esta rede foi utilizada então para previsões ao longo do período de 30-04-2014 a 25-06-2014. Note que

as saídas desta rede se encontram em valores normalizados, que para uso do segundo núcleo precisaram passar pelo processo inverso de normalização. Estas saídas ficam então na forma de uma previsão de preço, que são transformadas em previsão de retorno utilizando os valores de fechamento do dia anterior. Seguindo o exemplo, para o caso da ABEV3 no dia 30-04-2014, a saída do primeiro núcleo – após processo inverso de normalização – foi R\$ 14,57 sendo o valor de fechamento do dia anterior R\$ 14,37, ou seja, a previsão de variação de preço calculada foi de:

$$\frac{14,57 - 14,37}{14,37} = 1,39\%$$

Ao final do período é efetuado um novo treinamento da rede utilizando os dados de 25-07-2011 a 25-06-2014, e os pesos sinápticos obtidos são utilizados no período seguinte de 26-06-2014 a 19-08-2014. Assim seguiu com todos os ativos ao longo de todas as janelas.

Esta etapa foi repetida para cada um dos dez ativos, além do índice IBOV para construção do sistema de *Benchmark*. A tabela 7 apresenta uma síntese do desempenho médio destas redes calculadas com dados anteriores a 29/04/2014. As informações desta tabela foram posteriormente utilizadas como confiabilidade das previsões no modelo de Black-Litterman.

4.2 Segundo Núcleo: Modelo de Black-Litterman

O objetivo do segundo núcleo do sistema é receber as informações do primeiro núcleo e, juntamente com as informações da carteira Ibovespa diária, propor uma nova carteira diária utilizando posições de compra e venda para os dez ativos analisados, que correspondem aos dez ativos mais negociados no período do experimento. Devido à simplificação de redução do universo de ativos utilizados, o núcleo do otimizador de carteiras utilizou a composição diária do Ibovespa ajustada – calculado utilizando a equação (9) – para calcular o retorno excedente de equilíbrio diário Π_i . Alternativamente, como em Xing et al., 2018, o retorno de equilíbrio poderia ser calculado utilizando-se o método CAPM.

$$w_{i_{ajustado}} = \frac{w_i}{\sum_{10+} w_i} \quad (9)$$

O conjunto de previsões do primeiro núcleo, para cada dia, foi agrupado no vetor Q_i de visões dos investidores. Este vetor é composto pela saída das dez redes neurais transformadas em retornos esperados após processo inverso de normalização, como demonstrado ao final da seção 4.1.5.

Como estes retornos esperados são considerados não correlacionadas, a matriz P_i torna-se a matriz identidade. Foi utilizada uma matriz de covariância constante para todas os períodos do otimizador, calculada utilizando os dados dos conjuntos de treinamento e teste.

O retorno excedente de equilíbrio diário (Π_i) foi obtido através da otimização reversa do índice ajustado, utilizando a equação (10). Baseando-se em trabalhos anteriores, adotou-se δ , o coeficiente médio de aversão ao risco da carteira, fixado em 2.5. A variável w_i é a carteira diária do índice Bovespa e Σ a matriz de covariância.

$$\Pi_i = \delta \Sigma w_{i_{ajustado}} \quad (10)$$

Utilizando a abordagem de Idzorek, 2007, Walters, 2014 e Satchell and Scowcroft, 2000 para a implantação do modelo alternativo de Black-Litterman, foi definido $\tau = 1$ com o controle da combinação entre os retornos de equilíbrio e as estimativas das redes feita através da matriz Ω . Assim, para o cálculo dos retornos diários de Black-Litterman – com $\tau = 1$ e P_i igualando-se à matriz identidade – a equação (4) pode ser simplificada para a equação (11).

$$\hat{\Pi}_i = \Pi_i + \Sigma[\Sigma + \Omega]^{-1}[Q_i - \Pi_i] \quad (11)$$

4.2.1 Matriz Ω e confiabilidade α

Assim, a única variável ainda não definida é a matriz Ω . Como explicado anteriormente, na implantação do modelo alternativo pelo método de Idzorek, Ω é calculada utilizando uma abordagem de confiabilidade. Para as condições apresentadas, pode-se simplificar a equação (6) para a equação (12):

$$\begin{aligned} \Omega &= \alpha \text{diag}(\Sigma) \\ \alpha &= \frac{1 - C}{C} \end{aligned} \quad (12)$$

Onde Σ é a matriz de covariância da carteira – calculada utilizando valores históricos – e C é o vetor de confiabilidade composto pelo índice de acerto de operações das redes nos períodos de testes. O índice de acertos foi calculado considerando as operações vencedoras previstas pelas redes. Uma operação é considerada vencedora quando o retorno previsto pela rede (\hat{y}_s) possui o mesmo sinal do retorno real do mercado (y_s). Uma operação vencedora pode ser representada pela função passo de Heaviside, onde a operação j é vencedora quando $c_j = 1$ e perdedora se $c_j = 0$. Como descrito pela equação (13)

$$c_j = \frac{1 + \text{sgn}(s_j)}{2} \quad (13)$$

$$s_j = \text{sgn}(\hat{y}_s) * \text{sgn}(y_s)$$

A confiabilidade C_i da rede i pode então ser calculada contabilizando as operações vencedoras, ou pelo valor esperado de c_j , ou seja:

$$C_i = E[c_j] = \frac{\sum_j c_j}{j} \quad (14)$$

Os valores de confiabilidade das redes no período até 29/04/2014 estão expressos na tabela 7.

Ativo	Score
ABEV3	53.27%
BBDC4	51.30%
BRFS3	52.39%
CIEL3	51.08%
ITSA4	49.79%
ITUB4	51.63%
PETR3	50.35%
PETR4	51.81%
UGPA3	52.96%
VALE3	52.84%
IBOV	54.56%

Tabela 7: Confiabilidade das Redes

Como neste caso a variável α é um vetor, fica mais fácil analisar a equação (12) na forma da equação (5), ou seja:

$$\omega_{ij} = \alpha_i \Sigma_{ij}, \forall i = j$$

$$\omega_{ij} = 0, \forall i \neq j$$

onde :

$$\alpha_i = \frac{1 - C_i}{C_i} \quad (15)$$

Considere, por exemplo, o dia 02/05/2014. Os valores das variáveis do sistema para cada ativo podem ser observadas pela tabela 8.

	ABEV3	BBDC4	BRFS3	CIEL3	ITSA4	ITUB4	PETR3	PETR4	UGPA3	VALE3
$\hat{y}_s(\text{R}\$)$	14,57	17,76	48,23	20,30	5,18	22,77	16,16	16,76	50,72	25,76
Q_i	1,40%	1,84%	0,04%	-0,79%	0,40%	-0,39%	3,55%	1,34%	-0,09%	0,12%
Π_i	15,39%	5,49%	11,91%	6,77%	16,40%	4,91%	12,63%	12,95%	12,08%	6,49%
$\hat{\Pi}_i$	1,52%	2,67%	1,17%	0,42%	2,45%	2,42%	4,63%	4,28%	0,84%	2,71%
w'_{i+1}	7,19%	8,45%	-1,17%	-3,27%	-3,36%	1,06%	5,41%	6,71%	-4,43%	1,90%

Tabela 8: Estados do sistema para o dia 30/04/2014

A matriz Ω , tal qual a matriz Σ , são constantes nesta implementação. Utilizando Ω calculada a partir da tabela 7 e da equação (15), e substituindo os valores de Π_i e Q_i do dia 30/04/2014 na equação (11), chega-se aos retornos de Black-Litterman $\hat{\Pi}_i$, saída do segundo núcleo.

$$w'_i = (\delta \Sigma)^{-1} \hat{\Pi}_i \quad (16)$$

Os retornos de Black-Litterman, por sua vez, são utilizados então na equação de Markowitz sem restrições, equação (17), onde obtém-se a carteira w'_{i+1} para o dia 02/05/2014.

4.3 Sistema de Benchmark

O sistema de *Benchmark* foi construído utilizando as mesmas abordagens de construção do primeiro núcleo, tratando-se de uma implementação baseada no sistema de Gambogi and Costa, 2013. Ao contrário do modelo proposto neste trabalho, este

sistema de *Benchmark* monitora e opera apenas um ativo. A rede neural utilizada analisa as variações do índice Bovespa e se posiciona comprando ou vendendo o índice. Embora o índice não seja efetivamente um ativo negociado em bolsa, existem alguns ativos que seguem a variação do índice, como ETFs (*Exchange Trading Funds*).

A escolha do *Benchmark* pautou-se em criar um ambiente equivalente porém sem a utilização da inferência Bayesiana, representada pelo modelo de Black-Litterman. Após aplicação do mesmo processo seletivo do primeiro núcleo, a arquitetura III da tabela 1 foi a vencedora para o índice Bovespa. Utilizou-se também a técnica de janelamento móvel, com as mesmas dimensões de janela do sistema de dois núcleos.

O sistema de *Benchmark* utiliza as previsões de d_{i+1} fornecidas por uma rede neural, retreinada a cada início de janela de verificação, para se posicionar em relação ao índice Bovespa. Deste modo, quando a rede neural prevê uma oscilação positiva o investidor assume uma posição de compra do índice e quando a rede prevê uma oscilação negativa, assume-se uma posição de venda do índice. Ao final de cada dia de pregão, liquida-se a posição e repete-se a previsão do índice para o dia seguinte.

Tal qual no modelo de dois núcleos, o posicionamento do investidor é ponderado pela confiabilidade da rede neural durante o período de treinamento, com o Score médio da tabela 7, sendo que naquele esta confiabilidade alimenta o cálculo da matriz Ω , e neste o posicionamento é apenas ponderado de maneira linear.

Nestas condições, a valorização diária do patrimônio produzida pelo sistema de *Benchmark* é dada por:

$$V_{i+1} = s_i |p_{i+1} - p_i|$$

onde:

$$s_i = \begin{cases} +1 & \text{se } \text{ sinal}(p_{i+1} - p_i) = \text{ sinal}(y_{i+1} - p_i) \\ -1 & \text{se } \text{ sinal}(p_{i+1} - p_i) \neq \text{ sinal}(y_{i+1} - p_i) \end{cases}$$

onde p_i é o preço de fechamento no dia (d_i), p_{i+1} o preço de fechamento em (d_{i+1}), e y_{i+1} a previsão de preço de fechamento para (d_{i+1}).

4.4 Otimização de Carteiras

Após processamento de saída das redes, com as previsões alimentando o modelo de Black Litterman e os retornos a posteriori obtidos, o processamento dos dados do segundo módulo pôde ser feito pelo método de otimização de carteiras. A abordagem recomendada para os retornos de Black-Litterman é a otimização de Markowitz sem restrições, para o cálculo de carteiras diárias. Na equação (17), a carteira w'_i representa a posição a ser tomada utilizando os retornos Π_i obtidos na saída do segundo núcleo, para cada dia i .

$$w'_i = (\delta \Sigma)^{-1} \hat{\Pi}_i \quad (17)$$

As saídas da otimização sem restrição são analisadas no capítulo seguinte.

5 Resultados

As carteiras obtidas pelo modelo, em sua maioria, utilizam a estratégia *Long and Short*, se expondo negativamente a alguns ativos (posição vendida) e positivamente a outros (posição comprada). Desta maneira não seria incomum se o modelo acabasse negociando uma quantidade de ativos acima do seu patrimônio, ficando alavancado. Porém, para a aversão ao risco utilizada ($\delta = 2.5$), e para a metodologia utilizada na composição da matriz Ω , as carteiras produzidas se mostraram bastante conservadoras, com uma exposição máxima ao mercado de 72% em 05/12/2014. A exposição mínima ao mercado foi de 19% em 14/07/2015, e a exposição média foi de 35%.

Tanto no modelo proposto, quanto no modelo de *Benchmark*, o excedente do dinheiro não aplicado no mercado foi alocado no ativo livre de risco. A taxa de rendimento de depósito interbancário (CDI), índice de boa rentabilidade no período onde o modelo foi testado, foi escolhida como indexador do ativo livre de risco. Foi utilizada a taxa do CDI exercida ao longo do período de 30/04/2014 à 29/04/2016, mesmo período utilizado na verificação das redes neurais. Em relação à posição vendida, onde o investidor adota uma posição 'negativa', adotou-se prática comumente utilizada em corretoras brasileiras. Para uma posição vendida onde o investidor possui valor em caixa, a quantia investida é tomada como garantia pela corretora até a liquidação da operação, e caso o investidor não possua o valor em caixa a quantia investida é tomada como empréstimo sujeito a juros diários, com a posição sendo considerada alavancada. Porém, em nenhum momento os sistemas adotaram posições de alavancagem. A figura 8 mostra a frequência em dias dos índices de exposição ao mercado produzidos pelo modelo.

A figura 9 apresenta a evolução do valor das carteiras produzidas pelo modelo em comparação à evolução do valor produzido pelo sistema de *Benchmark*, com o índice Bovespa no mesmo período para referência.

A comparação dos desempenhos sob a ótica de risco e retorno pode ser feita pelas métricas da Tabela 9:

Nesta tabela o retorno anualizado é calculado obtendo o retorno total no período e transformando a taxa obtida para uma taxa de 12 meses. A volatilidade (risco) é calculada pelo desvio padrão dos retornos diários do sistema. O índice de Sharpe e o retorno excedente do investimento — subtraído do retorno do ativo livre de risco no período analisado, o índice DI — por unidade de

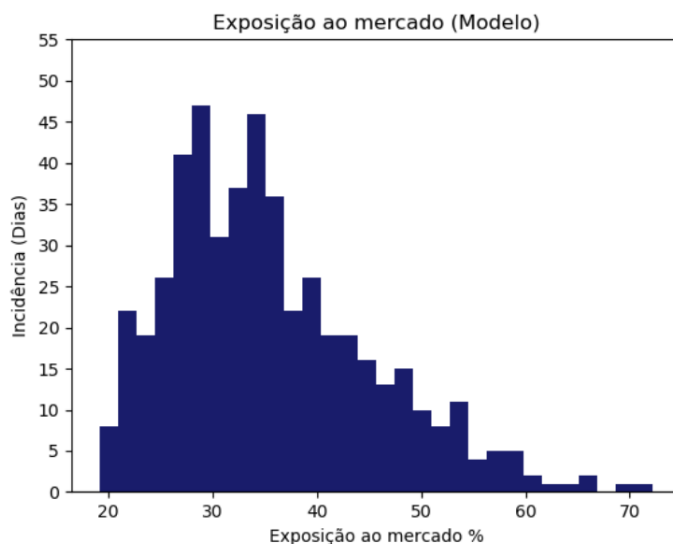


Figura 8: Exposição ao mercado

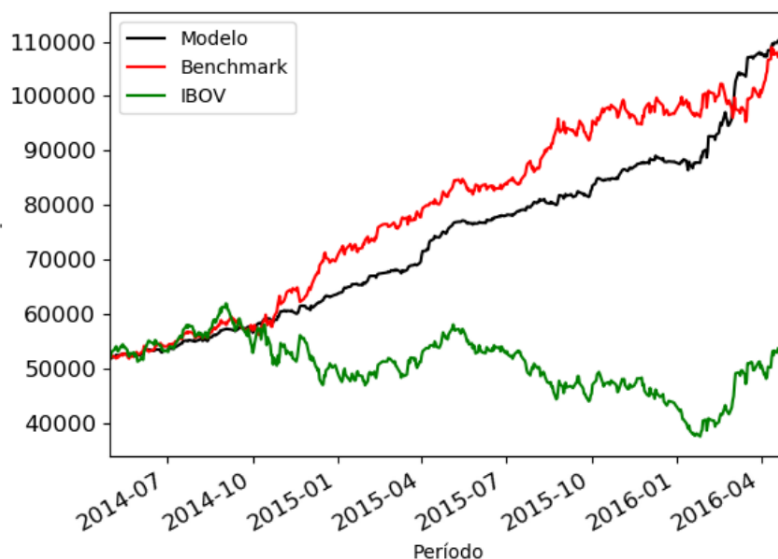


Figura 9: Curva de Evolução de Patrimônio

risco assumida, ou seja, dividido pelo desvio padrão dos retornos produzidos pela carteira. Ou seja:

$$IS = \frac{(r - r_f)}{\sigma}$$

O DrawDown máximo mede a variação negativa do maior vale observado na curva de retornos, calculado a partir do maior pico anterior ao início da queda.

Assim, quanto maior o índice de Sharpe, mais otimizada por unidade de risco encontra-se a carteira de ativos. O DrawDown é utilizado como um indicador comportamental e de risco, medindo o conforto do investidor e indicando qual o maior prejuízo da estratégia no período, antes dela voltar a ter um comportamento com ganhos. Assim, quanto menor o DrawDown, melhor o comportamento da carteira.

6 Conclusão

Tanto o sistema proposto quanto o *Benchmark* apresentaram ótima performance em relação ao Ibovespa. Porém, o sistema proposto apresentou um comportamento bastante superior ao *Benchmark* em relação à volatilidade, apresentando um índice de Sharpe com o dobro do valor e um DrawDown Máximo significativamente menor. Estes resultados fornecem argumentos a favor da abordagem utilizando inferência Bayesiana, e se mostrando uma ferramenta interessante no auxílio à gestão de carteiras. A principal vantagem da carteira que usa a estratégia do *Benchmark* em relação ao modelo proposto se deve ao fato da menor

	Retorno Anualizado	Volatilidade	Índice de Sharpe	DrawDown Máximo
Modelo	47.05%	07.73%	4.45	5.13%
Benchmark	43.96 %	14.24%	2.20	13.53%
IBOV	1.98%	26.35%	-0.41	47.07%

Tabela 9: Métricas de Desempenho

quantidade de operações necessárias, diminuindo o custo operacional.

Bons resultados podem ser obtidos em outras escalas de previsões, como as escalas semanais de Gambogi and Costa, 2013, o que reduziria os custos operacionais e aumentaria a viabilidade de implantação prática. Outra limitação do modelo é a metodologia de busca exaustiva em si, que demandou um enorme esforço computacional, mesmo com as simplificações adotadas. Outras abordagens, como a utilização de algoritmos genéticos, poderiam produzir os mesmos resultados com menor tempo de execução. O resultado mais interessante, entretanto, envolve as taxas de acerto das redes neurais. Nenhuma taxa de acerto individual das redes se manteve acima de 60%, resultado não muito impressionante. Porém, quando utilizadas em conjunto com a inferência bayesiana do modelo de Black-Litterman, apresentaram um ótimo retorno anualizado com baixa volatilidade, refletidos no índice de Sharpe. Foi observado que as taxas de acerto de muitas redes aumentaram em determinadas janelas no período de teste. Poderia-se aproveitar este comportamento atualizando-se a confiabilidade da equação (15) de maneira adaptativa, produzindo assim uma matriz Ω variável.

Outras abordagens em relação ao treinamento das redes, escolha de entradas e arquiteturas podem melhorar a taxa de acerto das redes, aumentando o retorno final do sistema.

Referências

- [Black and Litterman, 1992] Black, F. and Litterman, R. (1992). Global portfolio optimization. *Financial analysts journal*, 48(5):28–43.
- [Black and Litterman, 1991] Black, F. and Litterman, R. B. (1991). Asset allocation: combining investor views with market equilibrium. *The Journal of Fixed Income*, 1(2):7–18.
- [Chollet et al., 2015] Chollet, F. et al. (2015). Keras. <https://keras.io>.
- [Fabozzi et al., 2006] Fabozzi, F. J., Focardi, S. M., and Kolm, P. N. (2006). Incorporating trading strategies in the black-litterman framework. *The Journal of Trading*, 1(2):28–37.
- [Gambogi, 2013] Gambogi, J. A. (2013). Aplicação de redes neurais na tomada de decisão no mercado de ações. Master's thesis, Universidade de São Paulo.
- [Gambogi and Costa, 2013] Gambogi, J. A. and Costa, O. L. d. V. (2013). Sistema de trading no mercado de ações implementado por redes neurais artificiais. *Learning & Nonlinear Models*, 11(2):92–102.
- [He and Litterman, 2002] He, G. and Litterman, R. (2002). The intuition behind black-litterman model portfolios. *Investment Management Research*.
- [Idzorek, 2007] Idzorek, T. (2007). A step-by-step guide to the black-litterman model: Incorporating user-specified confidence levels. In *Forecasting expected returns in the financial markets*, pages 17–38. Elsevier.
- [Jones et al., 2001] Jones, E., Oliphant, T., Peterson, P., et al. (2001). SciPy: Open source scientific tools for Python. <http://www.scipy.org/>. [Online; accessed 26 november 2018].
- [Kaastra and Boyd, 1996] Kaastra, I. and Boyd, M. (1996). Designing a neural network for forecasting financial and economic time series. *Neurocomputing*, 10(3):215–236.
- [Kimoto et al., 1990] Kimoto, T., Asakawa, K., Yoda, M., and Takeoka, M. (1990). Stock market prediction system with modular neural networks. In *Neural Networks, 1990., 1990 IJCNN International Joint Conference on*, pages 1–6. IEEE.
- [Klassen, 2007] Klassen, M. (2007). Investigation of some technical indexes instock forecasting using neural networks. *International Journal of Computer, Electrical, Automation, Control and Information Engineering*, 1(5):1438–1442.
- [Lee, 2000] Lee, W. (2000). *Theory and methodology of tactical asset allocation*, volume 65. John Wiley & Sons.
- [Lo, 2004] Lo, A. W. (2004). The adaptive markets hypothesis. *The Journal of Portfolio Management*, 30(5):15–29.
- [Markowitz, 1952] Markowitz, H. (1952). Portfolio selection. *The journal of finance*, 7(1):77–91.

- [Meucci, 2006] Meucci, A. (2006). Beyond black-litterman in practice: A five-step recipe to input views on non-normal markets. *SSRN Electronic Journal*, 19.
- [Michaud, 1989] Michaud, R. O. (1989). The markowitz optimization enigma: Is 'optimized' optimal? *Financial Analysts Journal*, 45(1):31–42.
- [Murphy, 1999] Murphy, J. J. (1999). *Technical analysis of the financial markets: A comprehensive guide to trading methods and applications*. Penguin.
- [Oliphant, 2006] Oliphant, T. E. (2006). *A guide to NumPy*, volume 1. Trelgol Publishing USA.
- [Peters et al., 1994] Peters, E. E., Peters, E. R., and Peters, D. (1994). *Fractal market analysis: applying chaos theory to investment and economics*, volume 24. John Wiley & Sons.
- [Satchell and Scowcroft, 2000] Satchell, S. and Scowcroft, A. (2000). A demystification of the black–litterman model: Managing quantitative and traditional portfolio construction. *Journal of Asset Management*, 1(2):138–150.
- [Sharpe, 1964] Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, 19(3):425–442.
- [Walters, 2014] Walters, J. (2014). The black-litterman model in detail (june 20, 2014). Available at SSRN: <https://ssrn.com/abstract=1314585> or <http://dx.doi.org/10.2139/ssrn.1314585>.
- [White, 1988] White, H. (1988). Economic prediction using neural networks: The case of ibm daily stock return. In *Proceedings of International Conference on Neural Networks*, pages II–451.
- [Wutsqa and Kusumawati, 2016] Wutsqa, D. U. and Kusumawati, R. S. R. (2016). Radial basis function neural network for views prediction on black-litterman model. *Journal of Innovative Technology and Education*, 3(1):71–78.
- [Xing et al., 2018] Xing, F. Z., Cambria, E., and Welsch, R. E. (2018). Intelligent asset allocation via market sentiment views. *IEEE Computational Intelligence Magazine*, 13(4):25–34.
- [Zimmermann et al., 2001] Zimmermann, H. G., Neuneier, R., and Grothmann, R. (2001). Active portfolio-management based on error correction neural networks. In *Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic*, pages 1465–1472. MIT Press.