



ConBRepro

XIV CONGRESSO BRASILEIRO DE ENGENHARIA DE PRODUÇÃO

I CIGELUBRA

Congresso Internacional de Gestão e Engenharia Luso-Brasileiro

04 a 06 de dezembro de 2024

Energias Limpas nas Engenharias

Otimização de portfólio de investimentos de previdência complementar utilizando “machine learning” e “stacking”

Fernando Antônio Fernandes Garcia Pinto

Programa de Pós Graduação em Data Science

USP / Escola Superior de Agricultura "Luiz de Queiroz" da Universidade de São Paulo (Esalq)

Sérgio Mitihiro do Nascimento Maêda

Programa de Pós Graduação em Data Science

USP / Escola Superior de Agricultura "Luiz de Queiroz" da Universidade de São Paulo (Esalq)

Resumo: As Entidades de Previdência Complementar Fechada desempenham um papel crucial ao proporcionar aos participantes uma fonte segura de renda, principalmente para aposentadoria, sendo assim fundamental a gestão eficiente dos recursos financeiros para assegurar retornos satisfatórios. Este estudo utilizou a Teoria Moderna do Portfólio de Markowitz e técnicas de “machine learning” para otimizar a alocação de ativos de investimentos de um plano de previdência complementar, priorizando o equilíbrio entre risco e retorno e respeitando os limites de alocação estipulados. O objetivo visou explorar o impacto da utilização de uma combinação via “stacking” dos métodos de Redes Neurais Recorrentes, Prophet e SARIMA nas previsões de retornos de índices financeiros, com intuito de aumentar a qualidade do processo de otimização da alocação de ativos e gerar um inovativo e eficiente modelo de escolha de portfólio de investimentos para entidades de previdência complementar fechada. Os resultados mostraram que o “stacking” proporcionou previsões mais equilibradas e precisas, superando a meta atuarial estabelecida. O estudo evidencia a eficácia das técnicas de “stacking” e a importância da diversificação na gestão de ativos para planos de previdência.

Palavras-chave: Previdência complementar; Teoria Moderna do Portfólio; “Machine Learning”; Entidades de Previdência Complementar Fechada.

Optimization of Complementary Pension Investment Portfolio using Machine Learning and Stacking

Abstract: Pension Funds play a crucial role in providing participants with a secure source of income, especially for retirement, making the efficient management of financial resources essential to ensure satisfactory returns. This study utilized Markowitz's Modern Portfolio Theory and machine learning techniques to optimize the investment asset allocation of a pension plan, prioritizing the balance between risk and return while respecting the stipulated allocation limits. The objective was to explore the impact of using a “stacking” combination of Recurrent Neural Networks, Prophet, and SARIMA methods in predicting the returns of financial indexes, aiming to enhance the asset allocation optimization process and create an innovative and efficient investment portfolio selection model for pension funds. The results revealed that “stacking” provided more balanced and accurate forecasts,

surpassing the established actuarial target. The study highlights the effectiveness of "stacking" techniques and the importance of diversification in asset management for pension plans.

Keywords: Pension; Modern Portfolio Theory; Machine Learning; Pension Funds.

1. Introdução

A complementação de renda via previdência privada é essencial para manter o padrão de vida durante a aposentadoria, especialmente em países como o Brasil, onde o sistema previdenciário público enfrenta desafios significativos. As Entidades de Previdência Complementar Fechada [EFPC], de acordo com Coimbra e Toyoshima (2009), desempenham um papel crucial, fornecendo uma fonte segura de renda na aposentadoria. Garcia (2004) destaca que a gestão eficiente dos recursos financeiros dessas entidades é vital para garantir retornos satisfatórios, dado os compromissos de seguridade e a necessidade de maximização dos retornos dos ativos. Davis (1995) observa que tal gestão é fundamental para a preservação dos recursos previdenciários.

As EFPCs cresceram substancialmente nos últimos anos. Dados do relatório Consolidado Estatístico de outubro de 2023 da Associação Brasileira das Entidades Fechadas de Previdência Complementar [ABRAAP] indicam que os ativos dessas entidades somaram R\$ 1,22 trilhão, representando 11,2% do PIB, com mais de 8 milhões de participantes. A prioridade dos gestores das EFPCs é garantir rentabilidade adequada, cumprir as metas previdenciárias e satisfazer as expectativas dos participantes. A alocação de ativos é central para maximizar o retorno esperado para um nível de risco específico (Markowitz, 1952). Rubinstein (2002) aponta que prever eficazmente os retornos é fundamental, embora desafiador, dada a dinâmica do mercado financeiro. Métodos de otimização de portfólio baseados em aprendizado de máquina, conforme De Moura (2006), oferecem abordagens sofisticadas para lidar com essa complexidade.

A análise de séries temporais é uma abordagem comum na previsão de indicadores financeiros, utilizando dados históricos para identificar padrões. Modelos como SARIMA (Daryl et al., 2021) e Redes Neurais Recorrentes [RNR] (Selvin et al., 2017) têm demonstrado eficácia na modelagem sequencial dos preços das ações, permitindo previsões mais precisas. No entanto, a previsão financeira é complexa, já que fatores sociais, econômicos e políticos afetam os preços, tornando-os imprevisíveis. A hipótese do mercado eficiente, proposta por Cootner (1964) e Fama (1969), sugere que as variações de preços são aleatórias, limitando a eficácia das previsões financeiras. Contudo, outros estudos (Mandelbrot e Richard, 2004; Borges, 2010; Cova, 2011) desafiam essa teoria, apontando para assimetrias nos retornos e sugerindo a possibilidade de previsões, mesmo em cenários voláteis.

A proposta de Jiang et al. (2019) destaca a eficácia do "stacking" de diferentes modelos preditivos, sugerindo melhorias significativas na precisão das previsões ao combinar as forças individuais de cada modelo. Neste projeto, o "stacking" de RNN, Prophet e SARIMA é aplicado para aprimorar a gestão de ativos das EFPCs, criando um modelo inovador de otimização de portfólio.

2. Materiais e Métodos

2.1 Premissas Plano de Previdência Complementar Privada

Inicialmente, foi necessário criar um plano de previdência hipotético como base para o projeto. Definir as premissas do plano de previdência foi crucial para direcionar o estudo para a alocação ideal em um tipo específico de investimento institucional. Esse processo considerou planos de previdência complementar de referência no país, tanto em termos de

patrimônio quanto de participantes, de forma a modelar um plano hipotético que servisse de base para o estudo. O objetivo foi formular premissas que simulassem decisões de investimento e, assim, realizar um estudo que pudesse ser replicado em planos de previdência reais.

Para as definições das premissas, observaram-se três entidades de previdência complementar fechada de renome. As entidades e planos foram escolhidos diante de possuir informações de política de investimentos e características gerais de previdência de fácil e público acesso.

2.3 Modelo de previsão de retornos via “stacking”

O próximo passo foi a projeção de retornos futuros, utilizando três métodos distintos: RNN, Prophet e SARIMA.

As RNN, com sua capacidade de modelar relações complexas e não lineares, capturaram padrões ocultos nos dados de retorno. O modelo RNN utiliza a camada LSTM, que possui unidades de memória capazes de aprender e reter informações relevantes ao longo de períodos longos, superando limitações de redes neurais tradicionais. A LSTM foi configurada com 30 unidades, e o modelo foi treinado com dados históricos de retornos financeiros, utilizando a normalização dos dados e a preparação dos dados em janelas temporais. O modelo aprende os padrões sequenciais e, após o treinamento, realiza previsões futuras dos retornos financeiros. O Prophet, projetado para lidar com séries temporais que apresentam sazonalidade e tendências, foi utilizado para prever os retornos de forma robusta e automatizada. O modelo SARIMA, com suas características de captura de componentes sazonais e autorregressivos, foi aplicado para modelar o comportamento cíclico dos retornos.

Após a aplicação desses três métodos, utilizou-se o "stacking", um método de ensemble, que combinou as previsões geradas, proporcionando uma estimativa mais robusta e precisa. O stacking é uma técnica de ensemble learning que combina os resultados de diferentes modelos de aprendizado para gerar uma previsão mais robusta. Inicialmente, os três modelos são treinados e fazem previsões sobre o conjunto de dados de teste. Em seguida, essas previsões são ajustadas para que todos os modelos possuam o mesmo tamanho de saída. Novos dados são criados contendo as previsões de cada modelo, que são tratadas como variáveis independentes para um meta-modelo, neste caso, uma regressão linear. O meta-modelo é então ajustado para aprender a melhor maneira de combinar as previsões individuais e gerar uma previsão final, mais precisa. Esse processo também é aplicado às previsões futuras, onde o meta-modelo combina as previsões dos três modelos-base para fornecer estimativas para os próximos 12 meses. O objetivo do "stacking" é melhorar a precisão das previsões ao integrar as forças de cada modelo.

2.4 Otimização de portfólio de Markowitz e fronteira eficiente

O método de Markowitz (1952) fundamenta-se na escolha de carteiras que maximizem o retorno esperado para um dado nível de risco ou minimizem o risco para um retorno desejado. Utilizando as previsões geradas pelo "stacking", foi calculada a matriz de covariância dos retornos, essencial para identificar como os ativos se relacionam entre si e auxiliar na diversificação.

A otimização de portfólio foi realizada no Python, usando a abordagem de média-variância e o pacote SciPy para minimizar a variância do portfólio. Simulações de carteiras respeitaram os limites de alocação, buscando o melhor equilíbrio entre risco e retorno,

representado pela Fronteira Eficiente, que mostra carteiras com o maior retorno para cada nível de risco.

O Índice de Sharpe, utilizado para avaliar a eficiência das carteiras, foi calculado subtraindo a taxa de retorno livre de risco do retorno esperado da carteira e dividindo pelo desvio padrão dos retornos, indicando o retorno adicional por unidade de risco.

A matriz de covariância foi calculada com o Pandas, e o risco e o retorno esperados foram obtidos por operações matriciais no NumPy. O resultado final foi um portfólio otimizado que respeita os limites de alocação, maximizando a relação risco-retorno.

3. Resultados e Discussão

3.1 Premissas Plano de Previdência Complementar Privada

O plano de previdência complementar hipotético é composto por cinco segmentos de investimento: Renda Fixa (CDI e IMA-B), Renda Variável (Ibovespa), Exterior (IVVB11), Estruturado (IHFA) e Imobiliário (IFIX), com limites de alocação variando de 0% a 100% conforme o segmento. Os limites de alocação para cada segmento foram:

- a) Renda Fixa – mínimo 40%, máximo 100%;
- b) Renda Variável - mínimo 0%, máximo 30%;
- c) Exterior - mínimo 0%, máximo 20%;
- d) Estruturado - mínimo 0%, máximo 15%;
- e) Imobiliário - mínimo 0%, máximo 10%;

A meta atuarial é IPCA+4%, representando a rentabilidade necessária para cumprir os compromissos do plano. O estudo considerou um histórico de até 20 anos de dados para cada índice e desconsiderou observações atuariais e de passivo, como necessidade de liquidez e gestão de ativos e passivos, para focar exclusivamente na alocação do portfólio de investimentos.

A coleta e limpeza de dados incluíram índices como IPCA, IMA-B, CDI, IFIX, Ibovespa, IHFA e IVVB11, obtidos principalmente via API, exceto o IFIX, cujos dados foram acessados por planilhas disponíveis na internet. Os índices apresentaram diferentes históricos de dados, com o Ibovespa, CDI, IPCA, IHFA e IMA-B possuindo mais de 20 anos, enquanto o IFIX tem cerca de 13 anos de histórico. Essa variação não comprometeu a análise, pois todos os índices tinham dados suficientes para as previsões. As janelas mensais foram adotadas para o tratamento dos dados, unificando os retornos históricos mensais em uma única base de dados.

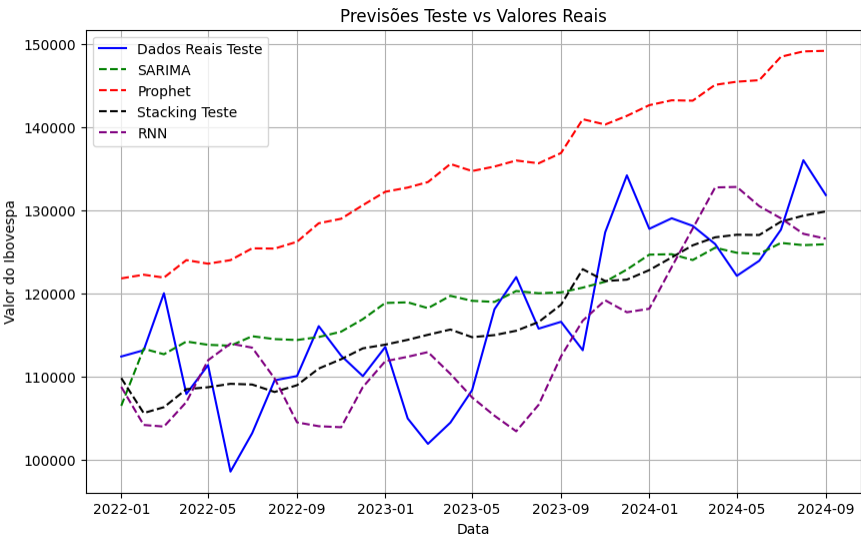
3.2 Modelo de previsão de retornos via “stacking”

De início, foram feitas projeções em separado para os índices de acordo com os três modelos de previsão propostos: Redes Neurais Recorrentes, Prophet e SARIMA. Cada índice em particular tem o seu histórico determinado, logo foi realizado uma separação em tamanhos diferentes para os “dataframes” de treinamento e de teste para cada modelo de previsão. Ao terminar de treinar cada um dos três modelos, foi realizada uma junção dos três modelos via “stacking”, baseado nos resultados adquiridos dos modelos com os dados de teste e realizado uma previsão combinada para os dados de teste.

Visualizando a Figura 1, resultado das previsões para os dados de teste do Ibovespa, o modelo Prophet, ao longo do período de previsão, mostrou um desvio significativo em relação aos dados reais, apresentando uma superestimação contínua, sem captar as flutuações e a sazonalidade observadas. Isso indica uma tendência de crescimento pouco realista, falhando em refletir as oscilações do mercado. Já o modelo RNN, conforme

discutido por Selvin et al. (2017), exibiu uma variabilidade acentuada, o que se alinha mais com o comportamento dos dados reais, mas ainda assim mostrou dificuldades em prever mudanças bruscas de tendência, especialmente durante períodos de volatilidade intensa no Ibovespa. Esses desafios refletem os limites das RNNs em captar flutuações extremas e padrões complexos.

Figura 1: Resultado das previsões dos dados de teste para o Ibovespa comparados aos valores reais



Fonte: O autor

O "stacking", que combina as previsões dos modelos Prophet, SARIMA e RNN, demonstrou ser uma abordagem mais equilibrada, tentando capturar os melhores aspectos de cada método: a suavidade do Prophet, a capacidade de detecção de sazonalidade do SARIMA e a habilidade das RNNs de lidar com oscilações mais pronunciadas. Mesmo assim, o "stacking" ainda não conseguiu reproduzir com precisão toda a complexidade e a volatilidade dos dados reais, embora tenha se aproximado mais deles em diversas ocasiões. O gráfico de desempenho sugere que o "stacking" é uma solução promissora para previsão de índices financeiros, buscando combinar o melhor de cada modelo, enquanto a Tabela 1 apresenta os indicadores estatísticos que reforçam as limitações e os pontos fortes de cada abordagem individual e combinada, especialmente em cenários de alta complexidade e volatilidade.

A Tabela 1 compara o coeficiente de determinação (R^2) dos diferentes modelos preditivos (Prophet, RNN e SARIMA) para uma série de índices financeiros, incluindo Ibovespa, IMA-B, CDI, IHFA, S&P500 e IFIX. O R^2 é uma métrica que indica o quão bem o modelo ajusta os dados, com valores que variam de 0 a 1. Um valor positivo próximo de 1 sugere um bom ajuste, enquanto valores negativos indicam que o modelo não consegue explicar adequadamente as variações observadas nos dados.

Tabela 1. Comparativo do desempenho dos modelos por indicadores estatísticos

Modelo	R^2	Índice
ARIMA	0.402160	Ibovespa
ARIMA	0.689364	CDI
ARIMA	-0.593359	S&P500
ARIMA	-0.018278	IMA-B
ARIMA	0.020500	IFIX
ARIMA	-0.837737	IHFA

Prophet	-2.825.708	Ibovespa
Prophet	0.408451	IMA-B
Prophet	-0.892151	CDI
Prophet	0.417801	S&P500
Prophet	-1.232.770	IFIX
Prophet	0.192140	IHFA
RNN	0.395225	IHFA
RNN	0.781659	S&P500
RNN	0.946626	CDI
RNN	0.350711	Ibovespa
RNN	-0.202777	IFIX
RNN	0.798849	IMA-B
Stacking	0.961390	IHFA
Stacking	0.912885	S&P500
Stacking	0.999684	CDI
Stacking	0.569280	Ibovespa
Stacking	0.722091	IFIX
Stacking	0.940193	IMA-B

Fonte: O autor

A previsão feita pelo Prophet apresentou dificuldades em capturar os padrões históricos dos índices IMA-B, Ibovespa e IHFA, com valores de R^2 negativos. Essa limitação foi semelhante à observada no estudo de Daryl et al. (2021), sugerindo que o Prophet pode não ser eficaz em contextos complexos e voláteis. Já o modelo RNN obteve os melhores resultados de R^2 para a maioria dos índices, destacando-se no S&P500 e IHFA, evidenciando sua capacidade de captar tendências não lineares. O SARIMA apresentou um desempenho misto, sendo mais preciso para o IMA-B e o S&P500, porém, menos eficaz para o CDI, IHFA e Ibovespa.

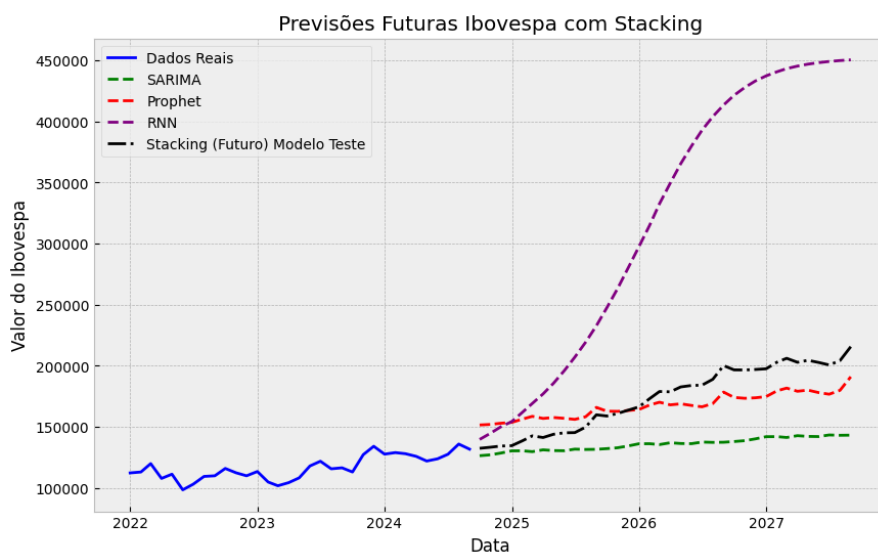
O “stacking”, que combina diferentes modelos, mostrou-se mais equilibrado, com R^2 positivos para a maioria dos índices, especialmente para o CDI e o IMA-B, mas ainda encontrou dificuldades com o IHFA e o IFIX. Em uma análise mais detalhada, a projeção do SARIMA indicou um crescimento conservador e estável para o Ibovespa, similar ao observado por Daryl et al. (2021) em estudos de ações da Apple. Isso reforça que, enquanto o SARIMA captura bem ambientes de baixa volatilidade, pode subestimar flutuações mais intensas.

A seleção de um modelo de previsão ideal depende das características específicas do índice analisado. Assim, a combinação de diferentes modelos, como feito no “stacking”, pode ser uma abordagem eficiente para melhorar a precisão preditiva. Essa análise pode ser útil para justificar a escolha mais apropriada para previsões financeiras e fornecer discernimento sobre a importância de avaliar diferentes técnicas preditivas para cada cenário.

Posteriormente, foi realizado também previsões individuais de cada modelo em separado, além do modelo “stacking”, para visualizar como seria o comportamento de previsão para 36 meses no futuro. Como presente na Figura 2, o modelo SARIMA apresentou previsões conservadoras, com o Ibovespa mantendo uma trajetória estável e crescimento modesto. Isso sugere que o SARIMA capturou padrões de baixa volatilidade e mudanças incrementais, o que pode ser útil em ambientes financeiros estáveis, mas subestima flutuações mais intensas ou crescimento acelerado. De forma semelhante, o trabalho de Daryl et al. (2021) analisou o desempenho do SARIMA para o mercado de ações da Apple,

destacando também uma tendência conservadora, mostrando que o modelo tende a suavizar movimentos bruscos em diferentes contextos. O modelo Prophet apresentou uma previsão moderada com flutuações cíclicas, sugerindo maior sensibilidade às sazonalidades do Ibovespa em comparação ao SARIMA, que mostrou uma abordagem mais conservadora e estável. Já o modelo RNN exibiu um crescimento exponencial nas previsões, destacando sua capacidade de capturar padrões complexos. No entanto, esse comportamento também pode indicar uma superestimação do crescimento futuro, como destacado por Selvin et al. (2017), que observou a superioridade das redes neurais em capturar dinâmicas não lineares em mercados financeiros.

Figura 2: Resultado das previsões dos modelos propostos e do “stacking” para os próximos 36 meses no futuro para o Ibovespa

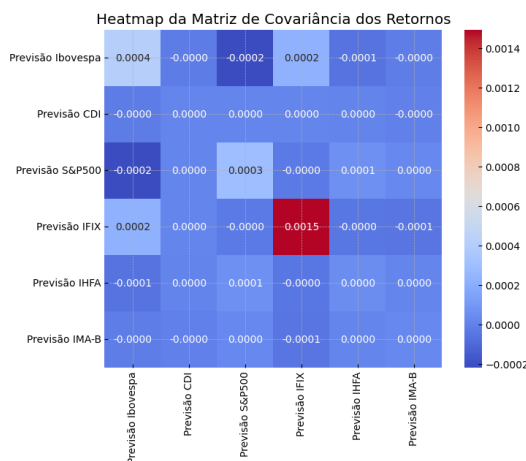


Fonte: O autor

3.3 Otimização de portfólio de Markowitz e fronteira eficiente

A matriz de covariância avalia como duas ou mais variáveis se movem em relação uma à outra, sendo crucial para entender a interação entre ativos financeiros.

Figura 3: “Heatmap” da Matriz de covariância dos retornos



Fonte: O autor

No “heatmap” da Figura 3, observa-se que covariâncias positivas indicam que os ativos tendem a se mover juntos, enquanto covariâncias negativas sugerem movimentos opostos. O gráfico revela que o IFIX possui maior volatilidade em relação a si mesmo e uma relação quase neutra ou negativa com outros ativos. O S&P500 (IVVB11) apresenta covariâncias muito baixas, sugerindo que é adequado para diversificação, enquanto o CDI tem covariâncias quase nulas, refletindo sua independência das variações de outros ativos. IHFA e IMA-B também mostram covariâncias baixas, indicando estabilidade. Com base nesses resultados, a alocação do IFIX foi reduzida de 15% para 5%, e a do IHFA, de 20% para 10%, visando minimizar o risco total do portfólio.

Em seguida, o problema de otimização foi formulado, onde o objetivo é encontrar a alocação ótima de pesos para os ativos que maximize o retorno ajustado pelo risco. Para garantir uma solução viável, foram aplicadas restrições, como a soma dos pesos igual a 1 (100% do capital alocado) e limites individuais de alocação para alguns ativos, assim como determinado nas premissas do projeto. Os resultados podem ser observados na Tabela 2 e 3.

Tabela 2. Resultado da otimização de limites de alocação

Segmento	Índice	Premissa de limite de alocação	Resultado otimização
Renda Variável	Ibovespa	mínimo 0%, máximo 30%	16.67%
Renda Fixa	CDI	mínimo 20%, máximo 100%	20,00%
Renda Fixa	IMA-B	mínimo 20%, máximo 100%	20,00%
Exterior	IVVB11 (S&P 500)	mínimo 0%, máximo 20%	16.67%
Imobiliário	IFIX	mínimo 0%, máximo 15%	5,00%
Estruturado	IHFA	mínimo 0%, máximo 10%	10.00%

Fonte: O autor

Tabela 3. Resultado do retorno e risco da carteira otimizada

Resultados	Valor
Retorno Médio Anual da Carteira Otimizada	15.62%
Retorno Médio Mensal da Carteira Otimizada	1.22%
Risco da Carteira Otimizada (Desvio Padrão):	0.38%

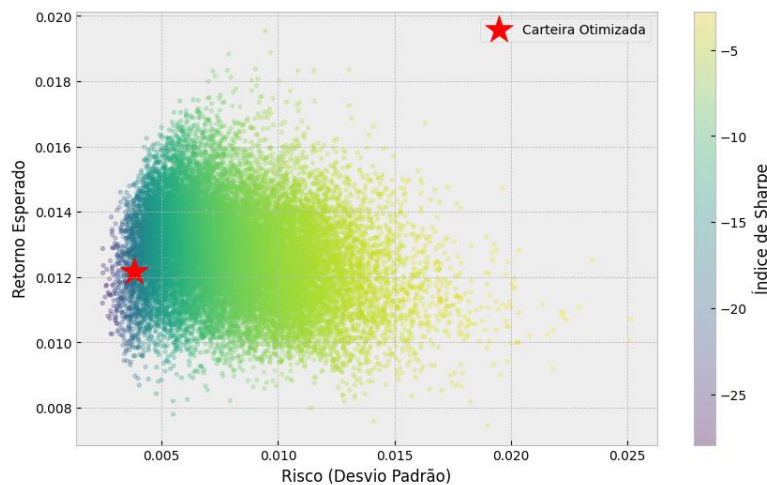
Fonte: O autor

A Tabela 2 detalha a alocação otimizada de ativos por segmento, respeitando os limites predefinidos para cada tipo de investimento e seguindo a metodologia de Markowitz, que visa maximizar o retorno esperado para um dado nível de risco. A alocação para o Ibovespa foi de 16,67%, mantendo-se dentro de uma faixa moderada e refletindo o potencial de retorno das ações brasileiras, sem elevar excessivamente o risco. CDI e IMA-B receberam 20% cada, garantindo segurança e previsibilidade, já que são ativos de renda fixa. Essa alocação sugere uma abordagem conservadora, dada a meta mínima de 20% para esses ativos. O S&P 500 (IVVB11) foi alocado em 16,67%, próximo ao limite de 20%, indicando uma visão otimista para o mercado norte-americano, proporcionando diversificação global ao portfólio.

O IFIX recebeu o mínimo de 5%, refletindo seu desempenho mais fraco nas projeções e a volatilidade associada. O IHFA foi alocado em 10%, atingindo o limite máximo, sugerindo que a diversificação por meio de fundos multimercados é considerada vantajosa, apesar de suas incertezas. A estratégia final demonstra uma preferência por ativos de renda fixa (CDI e IMA-B), combinada com exposição controlada a ações brasileiras e internacionais. A Fronteira Eficiente, representada na Figura 4, ilustra a relação risco-retorno das carteiras

simuladas, onde a carteira otimizada, destacada com uma estrela vermelha, se posiciona no extremo superior direito, indicando um alto retorno esperado com menor risco em relação à maioria das carteiras simuladas.

Figura 4: Fronteira Eficiente



Fonte: O autor

Os pontos espalhados no gráfico indicam diferentes combinações de ativos, com carteiras de baixo risco concentradas na parte inferior esquerda. A transição de cores (do roxo para o amarelo) indica o Índice de Sharpe, uma medida que ajusta o retorno ao risco. Carteiras com maiores valores de Índice de Sharpe são representadas pelas cores mais claras (amarelo), indicando um retorno mais alto para cada unidade de risco. O Índice de Sharpe é uma métrica de avaliação da relação entre retorno e risco. Um Índice de Sharpe mais alto indica que a carteira está gerando mais retorno por unidade de risco. No gráfico, as carteiras com índice de Sharpe mais alto tendem a estar mais próximas da fronteira eficiente e têm colorações que vão do verde ao amarelo claro.

O resultado do processo de otimização gerou os limites de alocação por segmento para a carteira do plano e o resultado do processo foi uma carteira balanceada, que prioriza a melhor alocação para que tenha o melhor retorno visando o menor risco possível. A carteira otimizada teve uma rentabilidade média de 15,62% ao ano. Dessa forma, superou os valores de meta atuarial projetados para os próximos três anos futuros. É evidente que a carteira otimizada não foi totalmente o ativo que trouxe maior rentabilidade em comparação com os índices estudados, como pode ser visualizado na Figura 8, mas foi a que mais trouxe adequação para as premissas fornecidas, principalmente visando uma alocação inteligente para um plano de previdência complementar fechado.

A Tabela 4 traz os resultados das projeções para os próximos 36 meses em formato de tabela, considerando a carteira otimizada.

O resumo da análise dos retornos projetados mostra que o S&P500 é o ativo com a maior rentabilidade esperada para os próximos 36 meses (132,13%), refletindo o otimismo com o mercado norte-americano. O Ibovespa também apresenta uma projeção expressiva de 63,64%, enquanto a carteira otimizada, com 54,53%, demonstra uma relação equilibrada entre risco e retorno, com foco na diversificação.

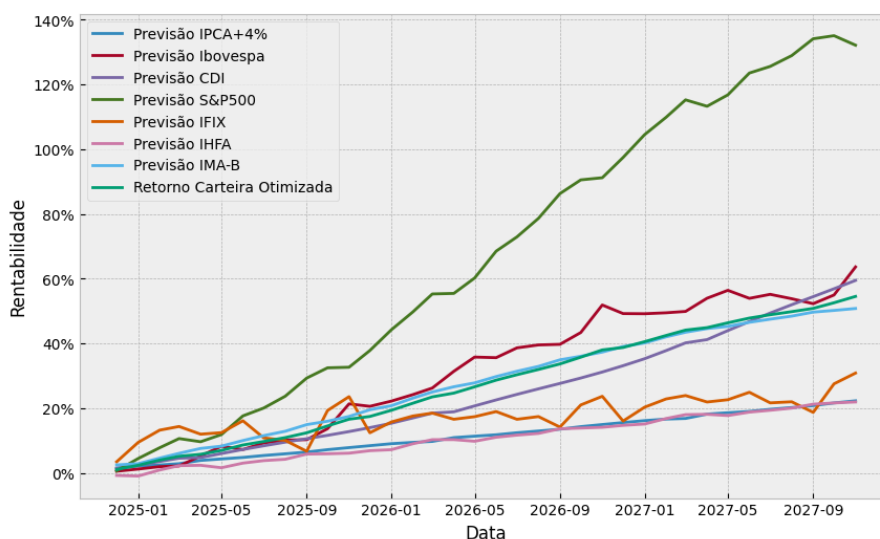
Tabela 4: Resultados de rentabilidade acumulada das projeções dos índices, da meta atuarial e carteira otimizada

Previsão	Rentabilidade prevista 36 meses	Risco (Desvio Padrão)
Previsão S&P500	132,13%	1,71%
Retorno Carteira Otimizada	64,06%	0,38%
Previsão Ibovespa	63,64%	2,03%
Previsão CDI	59,48%	0,40%
Previsão IMA-B	50,80%	0,48%
Previsão IFIX	30,85%	3,86%
Previsão IPCA+4%	22,28%	0,25%
Previsão IHFA	21,93%	0,66%

Fonte: O autor

Os retornos de CDI (59,48%) e IMA-B (50,80%) indicam que os ativos de renda fixa ainda podem proporcionar bons ganhos em um cenário de juros elevados. Por outro lado, o IFIX (30,85%) e o IHFA (21,93%) apresentam previsões mais modestas, refletindo maior cautela no setor imobiliário e nos fundos multimercados. Já o IPCA+4% projeta um retorno de 22,28%, demonstrando um arrefecimento na inflação dos próximos anos, o que facilita a carteira do plano a superar sua meta atuarial. A análise destaca a solidez da carteira otimizada, que oferece retornos competitivos com menor volatilidade em comparação a ativos de maior risco.

Figura 5: Rentabilidade acumulada das projeções dos índices, da meta atuarial e carteira otimizada



Fonte: O autor

4. Conclusão

Este trabalho apresentou uma abordagem inovadora para a otimização de carteiras em EFPCs ao integrar métodos avançados de ML com a Teoria Moderna do Portfólio de Markowitz. A metodologia aplicada demonstrou a relevância da combinação de modelos preditivos, como Prophet, SARIMA e RNN, via ensemble de "stacking", para a projeção dos retornos futuros dos ativos.

Através da análise dos dados históricos dos índices selecionados, foi possível observar o desempenho diferenciado de cada modelo preditivo, com destaque para o método de RNN que se mostrou mais robusto em alguns cenários. A técnica de "stacking" permitiu melhorar

a precisão das previsões ao combinar os pontos fortes de cada modelo, fornecendo resultados mais precisos para a otimização da alocação de ativos.

Na etapa de otimização de portfólio, a aplicação da fronteira eficiente de Markowitz proporcionou uma distribuição ponderada dos ativos que priorizou o equilíbrio entre risco e retorno, respeitando os limites de alocação predefinidos para o plano de previdência. A carteira otimizada conseguiu superar a meta atuarial de IPCA+4%, apresentando um retorno alto para um risco baixo em comparação com mais de 20 mil carteiras simuladas, o que reforça a viabilidade da abordagem proposta.

Assim, a aplicação de técnicas de ML em conjunto com a teoria de Markowitz mostrou-se uma estratégia eficaz para a otimização de portfólios em EFPCs, oferecendo uma solução prática e replicável para melhorar o desempenho das carteiras e alcançar metas previdenciárias. Esse trabalho contribui para o desenvolvimento de novas ferramentas de apoio à gestão financeira, demonstrando o potencial de uso de tecnologias emergentes no campo de investimentos e na previdência complementar.

Para futuros trabalhos, o acréscimo de novas informações e dados podem ser explorados. Utilizar dados econômicos (resultado das empresas, desemprego, serviços, câmbio, produção, juros, etc.) além do histórico dos índices podem fornecer reconhecimento de padrões ainda mais complexos para o modelo treino e, dessa forma, fornecer previsões ainda mais coerentes e conceituadas. Adicionalmente, a exploração de outros métodos de inteligência artificial para utilização no “stacking” ou outro método de “ensemble” podem fornecer também ainda mais complexidade ao tema.

5. Referências

Associação Brasileira das Entidades Fechadas de Previdência Complementar [ABRAAP]. Outubro de 2023. **Relatório Consolidado Estatístico**.

Breiman, L. 1996. **Bagging predictors**, Machine learning, vol. 24: 123–140.

Borges, R. 2010. **Eficiência e o Mercado de Renda Variável Brasileiro**. Dissertação (Mestrado) Universidade Cândido Mendes, Rio de Janeiro, Brasil.

Cova, C. J. G. 2011. **Finanças e mercado de capitais - mercados fractais**: A nova fronteira das finanças. São Paulo. p. 23–46.

Coimbra, C.B.; Toyoshima, M.T.K. 2009. **Uma análise do setor de previdência complementar brasileiro**. Revista de Economia Contemporânea, Rio de Janeiro: v. 13: 439–466.

Cootner, P. 1964. **The random character of stock market prices**. Cambridge: MIT Press.

Daryl, A. Winata, S. Kumara and D. Suhartono. 2021. "Predicting Stock Market Prices using Time Series SARIMA," 2021 **1st International Conference on Computer Science and Artificial Intelligence**. p. 92-99.

Davis, E.P. 1995. **Pension Funds—Retirement-Income Security, and Capital Markets—An International Perspective** (Oxford: Clarendon Press).

De Moura, G. 2006. **O uso de redes neurais artificiais na previsão de tendências no mercado de ações** (PhD Thesis).

Fama, E. F. 1969. **Efficient capital markets: A review of theory and empirical work**. The Journal of Finance, v. 2: p. 383–417

Gajamannage, K.; Kobayashi, M.; Miyakawa, S.; Takahara, Y. 2023. **Real-time forecasting of time series in financial markets using sequentially trained dual-LSTMs**.

- Garcia, M.M. 2004. An Analysis of Pension Funds in Portugal, Pensions: **An International Journal**, 9(3): 227-245.
- Garlapati, A., Krishna, D. R., Garlapati, K., Srikara Yaswanth, N. mani, Rahul, U., & Narayanan, G. 2021. **Stock Price Prediction Using Facebook Prophet and Arima Models**.
- Jiang, M., Liu, J., & Zhang, L. 2019. **An improved Stacking framework for stock index prediction by leveraging tree-based ensemble models and deep learning algorithms**.
- Lee, Cheng-Min; KO, Chia-Nan. 2011. Short-term load forecasting using lifting scheme and ARIMA models. **Expert Systems with Applications**. p. 5902-5911.
- Ma, Y.; Principe, J.C. 2018. "Comparison of Static Neural Network with External Memory and RNNs for Deterministic Context Free Language Learning," 2018 **International Joint Conference on Neural Networks (IJCNN)**. p. 1-7.
- Mandelbrot, B. B.; Richard, H. 2004. **Mercados financeiros fora de controle: A teoria dos fractais explicando o comportamento dos mercados**.
- Markowitz, H. 1952. Portfolio selection. **The Journal of Finance**, 7(1): 77-91.
- Rubinstein, M. 2002. Markowitzs Portfolio Selection: A Fifty-Year Retrospective. **The Journal of Finance** 57 : 1041-1045.
- Saud, Z.A.; Shakya, S. 2020. Analysis of look back period for stock price prediction with rnn variants: A case study on banking sector of nepse. **Procedia Computer Science**: v. 167: 788–798
- Selvin, S.; Vinayakumar, R.; Gopalakrishnan, E.A.; Menon, V.K.; Soman, K.P. 2017. Stock price prediction using LSTM, RNN and CNN-sliding window model. 2017 **International Conference on Advances in Computing, Communications and Informatics (ICACCI)**. p.1643-1647.
- Ryll, L.; Seidens, A. 2019. **Evaluating the Performance of Machine Learning Algorithms in Financial Market Forecasting: A Comprehensive Survey**.
- Rossi, J. W; Neves, C. 2014. **Econometria e Séries Temporais com Aplicações a Dados da Economia Brasileira**.
- Zaremba, W.; Sutskever, I.; Vinyals, O. 2014. **Recurrent neural network regularization**. p. 1409.2329.