

HERMES

Simulando investidores com IA para maximizar
retornos no mercado de criptoativos



GERÔNIMO PULITINI | HELENA METELLO | HUGO HICKMAN | ISABELLE LEAL | MAURÍCIO LIMA

INTRODUÇÃO

Sumário

1 Introdução

2 Metodologia

6 Dados

7 Resultados

8 Conclusão

9 GitHub

Motivação

A rápida expansão do mercado de criptomoedas e a sua alta volatilidade impõe desafios significativos para investidores e analistas financeiros, ao mesmo tempo em que a inteligência artificial tem possibilitado o desenvolvimento de sistemas capazes de aprender padrões complexos e prever movimentos de mercado com maior precisão.

Delimitação do problema

Este estudo pretende simular investidores veteranos e consolidados do mercado de criptomoedas através de um sistema multiagente de inteligência artificial generativa, a fim de analisar como diferentes perfis de tomada de decisão podem influenciar a maximização de retornos e a redução de riscos em um ambiente de alta volatilidade.

METODOLOGIA

Estratégia utilizada

A estratégia consiste na construção de uma arquitetura multiagente de inteligência artificial, em que cada agente desempenha uma função específica no fluxo operacional, acessando uma camada de RAG para buscar conhecimentos relevantes para suas atividades. Essa arquitetura simula o processo mental de um investidor profissional, reproduzindo sua tese desde a pesquisa das criptomoedas até a elaboração do relatório final com as informações relevantes e os resultados.

O investidor escolhido foi **Changpeng Zhao**, um dos nomes mais influentes do mercado cripto. Fundador da Binance, ele transformou a indústria ao combinar visão estratégica, tecnologia e foco na utilidade prática dos ativos digitais. Seu estilo de investimento valoriza fundamentos sólidos, ciclos de mercado, construção de comunidade e produtos com uso real. É exatamente esse pensamento que buscamos emular para construir uma carteira de um ano.

Para tal, estruturamos uma base abrangente de informações sobre Changpeng Zhao, incluindo entrevistas, histórico de decisões de compra e venda, análises públicas, pronunciamentos em períodos de alta e de crise, além de obras publicadas a seu respeito. Todo esse conteúdo foi fragmentado, vetorizado e organizado com RAG, que utiliza álgebra linear para encontrar apenas os trechos mais relevantes antes do envio ao modelo de linguagem. Assim, o LLM recebe diretamente seus princípios fundamentais, permitindo construir um clone sintético do investidor.

A solução envolve duas frentes de análise, aproveitando a capacidade das inteligências artificiais gerativas de consumir volumes elevados de informação e avaliar conteúdos em níveis semânticos e generalizados. A primeira frente é quantitativa e utiliza algoritmos de regressão para prever a distribuição das criptomoedas e gerar o Expected Return Score e métricas de risco usando simulações de Monte Carlo. A segunda frente é qualitativa e absorve informações de mercado, como postagens do Reddit, notícias e relatórios. A partir de técnicas de RAG, em que usamos um banco vetorial, a IA captura o estilo decisório e os padrões de julgamento do investidor selecionado, transformando sinais qualitativos em um Conviction Score mensurável, que representa o nível de confiança da IA em cada criptomoeda.

O Conviction Score, Expected Return e as métricas de risco são integrados em uma única carteira por meio do modelo matemático Black-Litterman, que concilia opinião qualitativa e dados quantitativos em uma estrutura otimizada de risco e retorno. A proposta amplia a capacidade de julgamento dos investidores humanos, aliando convicção qualitativa e optimização matemática em uma abordagem robusta para decisões de investimento.

METODOLOGIA

Modelagem Operacional

O modelo opera em ciclos mensais e segue um fluxo sistemático composto por cinco etapas principais:

1) Coleta de dados

A cada ciclo, são executadas duas rotinas paralelas que posteriormente convergem:

- Pipeline Quantitativo
- Pipeline Qualitativo

2) Seleção dos ativos

- Universo Inicial: Criptomoedas
- Filtros Eliminatórios: Seleção preliminar da IA baseado no investidor
- Carteira Final: Ativos escolhidos (peso otimizado)

3) Geração de sinais

Cada ação recebe dois scores independentes:

- Qualitativo:
 1. Conviction Score (0–10)
 2. IA generativa + RAG
- Quantitativo:
 1. Expected Return Score
 2. Modelagem estatística, regressões fatoriais

4) Construção da carteira

Os dois vetores de informação convergem no modelo Black-Litterman, em que:

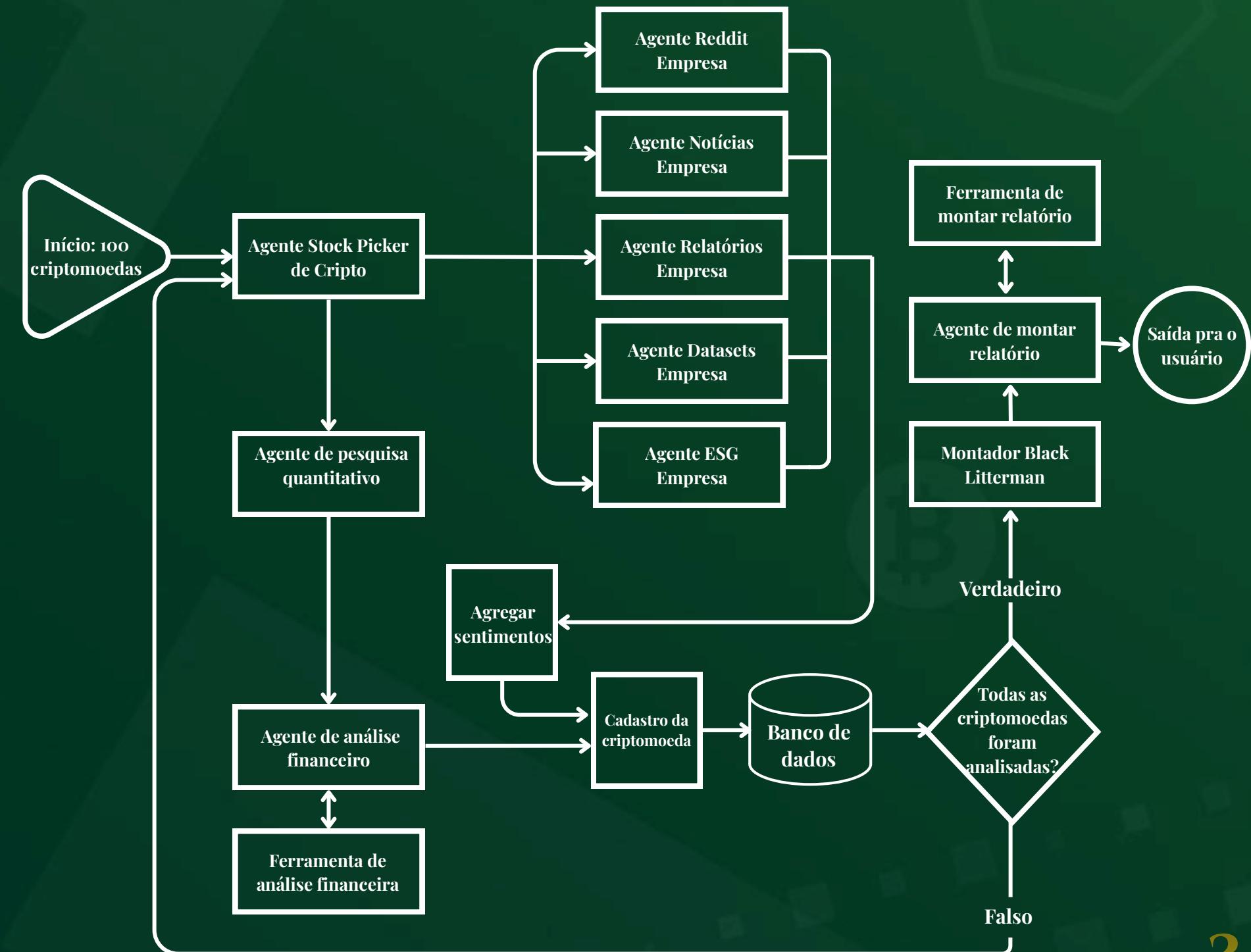
- Os agentes de IA quantitativos: fazem previsões e análises de risco
- Os agentes de IA qualitativos: fazem a análise de sentimentos, de modo a emular o investidor
- Resultado final: Pesos ótimos com restrição de apenas comprar ações

5) Modelagem de previsão e Restrições

- Frequência de previsão: Mensal
- Peso mínimo: Sem limite / peso máximo: 20%
- Sem alavancagem ou posições vendidas.

Fluxograma

Ajuda a entender toda a cadeia de processos, desde a formação da carteira até o resultado final



METODOLOGIA

Modelos preditivo

As famílias de modelos utilizadas incluem as Baselines (naive, moving average e rolling linear), os modelos ARIMA, SARIMA e SARIMAX, os modelos ETS (Error, Trend, Seasonality), os modelos de Estado-Espaço como o Kalman Local Linear Trend (LLT), os algoritmos de TreeBoosting (XGBoost e LightGBM), além de modelos auxiliares voltados ao tratamento de variáveis exógenas, como o ExogLSTM e o Seq2Seq, e um modelo distribucional (DistModel) voltado à estimação de distribuições preditivas completas.

Backtest

Treino

Nessa etapa é feito o diagnóstico da série temporal (tendência, sazonalidade, covariância a 95%), os dados foram utilizados para preparar os modelos preditivos, exógenas foram filtradas e foram definidas famílias de modelos compatíveis.

Teste

Os hiperparâmetros foram otimizados com Optuna, incluindo os ligados à construção do dataset. Após comparar os candidatos pela métrica-alvo, selecionou-se o melhor tipo de modelo e seus hiperparâmetros foram refinados em um ajuste final.



Avaliação

Permanência do modelo fixo, aplicação time-lock para informações qualitativas e aplicado o backtest da carteira; foi reportado performance vs. benchmark e métricas (retorno, Sharpe, desvio padrão, dropdown, alpha etc.), tendo as mesmas janelas/regras para todas as criptomoedas, sem look-ahead bias e nem reotimização.

DADOS

Modelagem quantitativa de retorno de risco

Beta provenientes

Retornos

Volume

Preços



Extração qualitativa via LLM

Uso de Gemini pela sua capacidade de utilizar 1 milhão de tokens (700 páginas).



Sensibilidade Macro para views

Dados macroeconômicos



Proxy de narrativa e hype

Redes sociais -
API do Reddit



Eventos e notícias gerais

Notícias



Notícias
financeiras



RESULTADOS

Análise de risco

A carteira apresenta retorno anualizado superior ao do HASH11 (54,9% versus 46,3%), contudo esse resultado é obtido à custa de um nível de risco significativamente mais elevado: a volatilidade anualizada é substancialmente maior, o máximo drawdown aproxima-se de -49% e o índice de Sharpe permanece inferior ao do ETF, indicando que o prêmio de retorno não compensa integralmente o aumento das oscilações. Por outro lado, o alfa anualizado positivo, em torno de 17%, e o Treynor Ratio superior ao do benchmark sugerem que há contribuição efetiva da seleção de ativos, e que a carteira gera retorno excedente em relação ao risco sistemático assumido. O VaR diário mais negativo e o Hit Ratio apenas marginalmente melhor que o do HASH11 reforçam a conclusão central: trata-se de uma estratégia de perfil marcadamente agressivo.

MÉTRICA	CARTEIRA	HASH11
Retorno Anualizado	54,91%	46,33%
Volatilidade Anualizada	66,58%	45,74%
Volatilidade Acumulada	54,13%	45,68%
Sharpe Ratio	0,765	0,925
Max Drawdown	-49,01%	-37,79%
VaR 95% (Diário)	-5,69%	-4,12%
Alfa Anualizado (CAPM)	16,99%	0%
Treynor Ratio	0,510	0,423
Hit Ratio (% dias positivos)	52,53%	51,36%



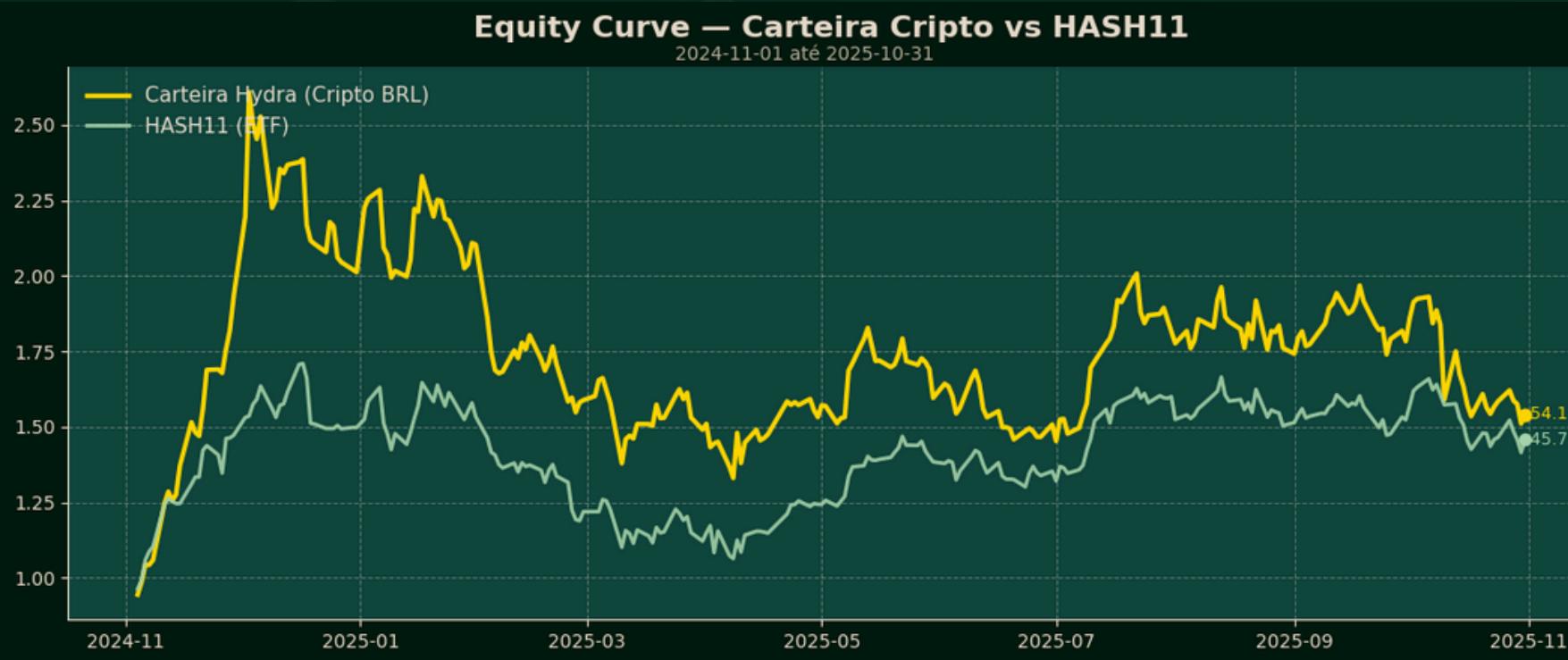
Disposição da carteira

A carteira é concentrada em altcoins e respeita um hardcap de 20% por ativo: TRX e XRP encostam nesse limite com 20% cada, seguidas por BTC (15%) e DOT (12,5%), formando um núcleo de 67,5% em posições de alta convicção. As maiores alocações foram direcionadas para moedas com tese fundamental clara e valor intrínseco (pagamentos, infraestrutura, contratos inteligentes), e não para memecoins, refletindo o papel da análise qualitativa na seleção dos vencedores e no uso “cheio” do hardcap apenas onde o modelo enxerga melhor relação risco–retorno. Na faixa intermediária, NEAR (10,1%) e SOL (9,6%) reforçam a aposta em plataformas de alta performance, enquanto a cauda com TON (5%), AVAX (4,9%) e SHIB (3%) adiciona diversificação mais especulativa. No conjunto, é uma carteira agressiva, bem diversificada em nomes, porém com baixa representatividade de instrumentos de caráter mais defensivo.

RESULTADOS

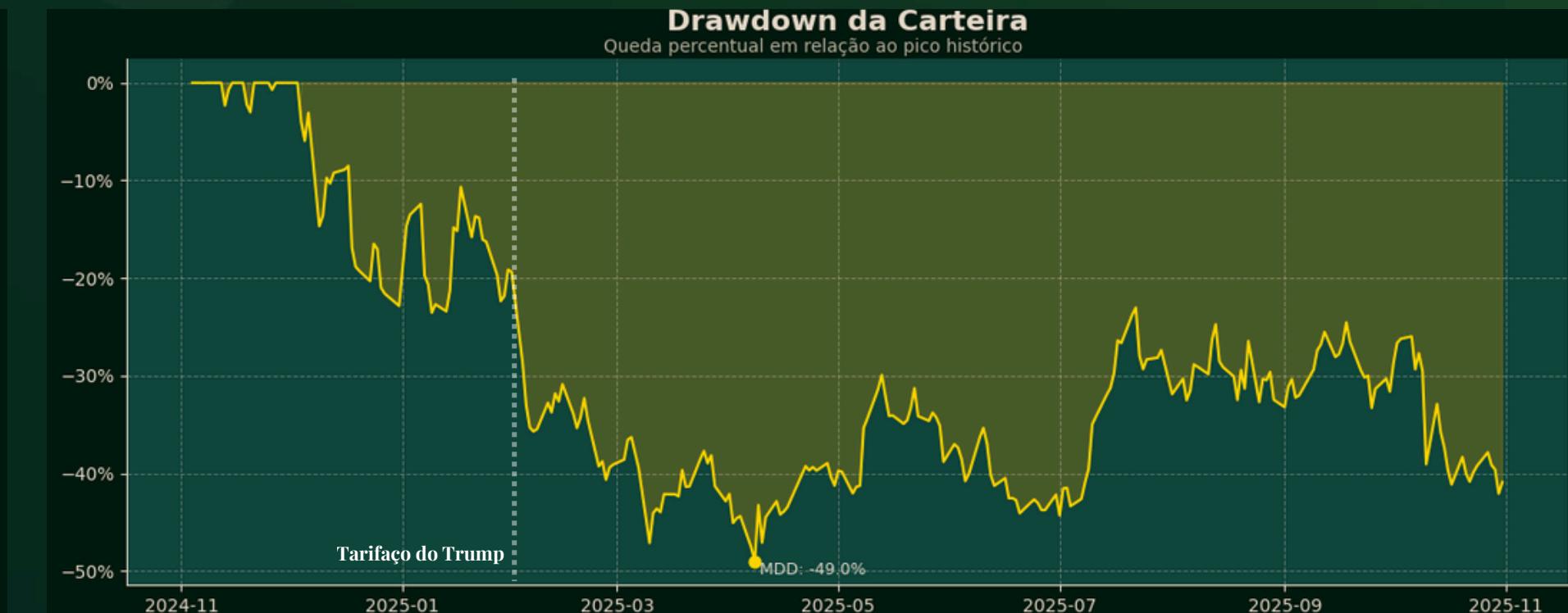
Equity Curve

A estratégia começa praticamente pareada com o HASH11, mas se destaca no forte rali do fim de 2024, disparando bem à frente do ETF; mesmo após a correção acentuada no início de 2025, recupera-se mais rápido e encerra o período mantendo uma vantagem consistente sobre o benchmark (menor em relação ao início).



Drawdown

O drawdown aprofunda logo após o tarifaço nos EUA: o mercado já apresentava um ambiente de forte euforia no final de 2024 e início de 2025, caracterizado por posições altamente alavancadas. Nesse contexto, o anúncio das tarifas desencadeou um movimento global de redução de risco (risk-off); essa combinação de alavancagem elevada + fuga de risco faz a carteira cair bem mais forte que o rali anterior, chegando perto de -50% antes de uma recuperação apenas parcial.



CONCLUSÃO

Síntese dos Resultados e Relevância da Metodologia

O estudo mostra que a combinação de análises quantitativas, modelos preditivos e IA generativa em um sistema multiagente melhora significativamente a qualidade das decisões no mercado de criptomoedas, marcado por alta volatilidade, assimetria de informações e forte influência de narrativas.

A integração de previsões estatísticas, sinais de séries temporais, métricas de risco e um Conviction Score baseado em notícias, redes sociais e narrativas amplia a compreensão dos fatores que movem as criptomoedas. Esses vetores de informação são consolidados em uma estrutura de otimização que identifica, a cada ciclo, o ativo com benefícios não capturados por modelos convencionais, replicando o processo decisório de um investidor especializado.

Conclui-se que a integração entre IA generativa, modelagem quantitativa e análise de narrativas configura uma estratégia sólida para recomendações de investimento em criptomoedas. O sistema produz recomendações consistentes e ajustadas às particularidades do mercado, tornando-se uma ferramenta promissora para decisões estratégicas em ambientes de alta incerteza.

Os resultados indicam que a abordagem híbrida aumenta a precisão das previsões, reduz ruídos típicos da alta volatilidade e reage com eficiência a choques repentinos de mercado. A metodologia identifica oportunidades de retorno assimétrico enquanto controla a exposição a perdas acentuadas, um requisito fundamental no mercado de criptomoedas.

Próximos Passos

Realizar novos testes de robustez:

Validar o desempenho da metodologia em diferentes condições de mercado.

Ampliar gradualmente o universo de criptomoedas analisadas:

Incluir novas moedas líquidas para melhorar a comparação entre opções.

Refinar o Conviction Score:

Aperfeiçoar o peso das fontes de informação para tornar o sinal qualitativo mais preciso.

Implementar rebalanceamento semanal:

Ajustar a frequência da estratégia para reduzir risco e suavizar variações extremas.

LINK PARA GITHUB

<https://github.com/geronimopulitini/Desafio-FGV-Datathon-Equipe-Hermes>



TIME ENVOLVIDO

Gerônimo Pulitini



Ciências
Econômicas
PUC - RIO

Helena Metello



Engenharia de
Produção
PUC - RIO

Hugo Hickman



Engenharia de
Produção
PUC - RIO

Isabelle Leal



Estudos de
Mídia
PUC - RIO

Maurício Lima



Engenharia de
Produção
PUC - RIO