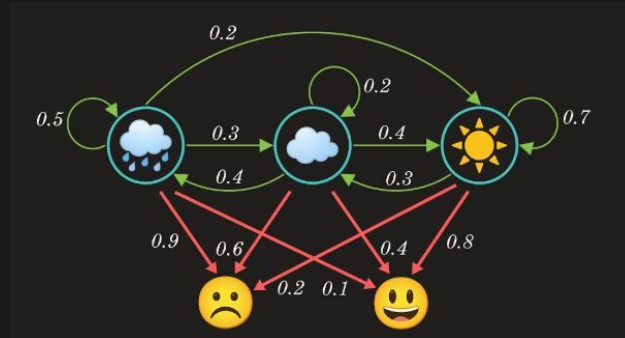




# Showcase Diretoria de Tecnologia

## *Deep markovian trader*



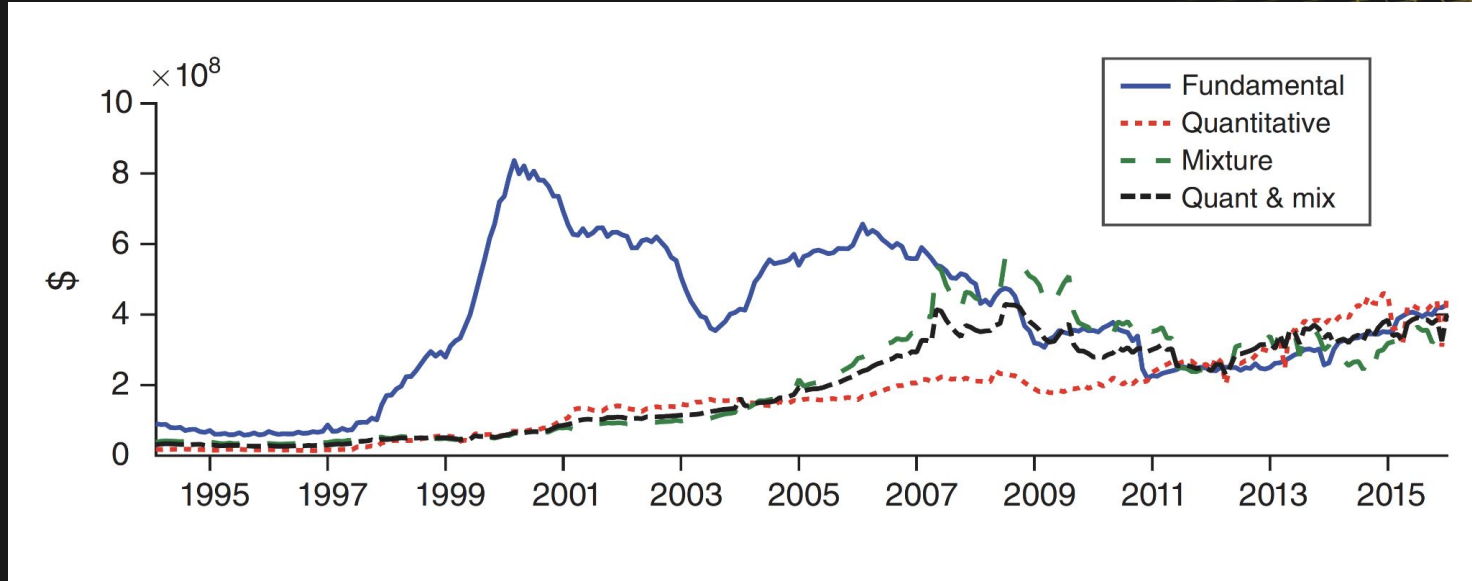


# Objetivos da Apresentação

- Introduzir o **mercado de finanças quantitativas**, destacando seu potencial e oportunidades.
- Apresentar a **construção de uma estratégia quantitativa**, desde a filtragem de ativos até a geração de sinais de trading.
- Explicar como **modelos econométricos** (ex: MSGARCH) e **redes neurais** (ex: LSTM) podem ser combinados para prever o comportamento do mercado.

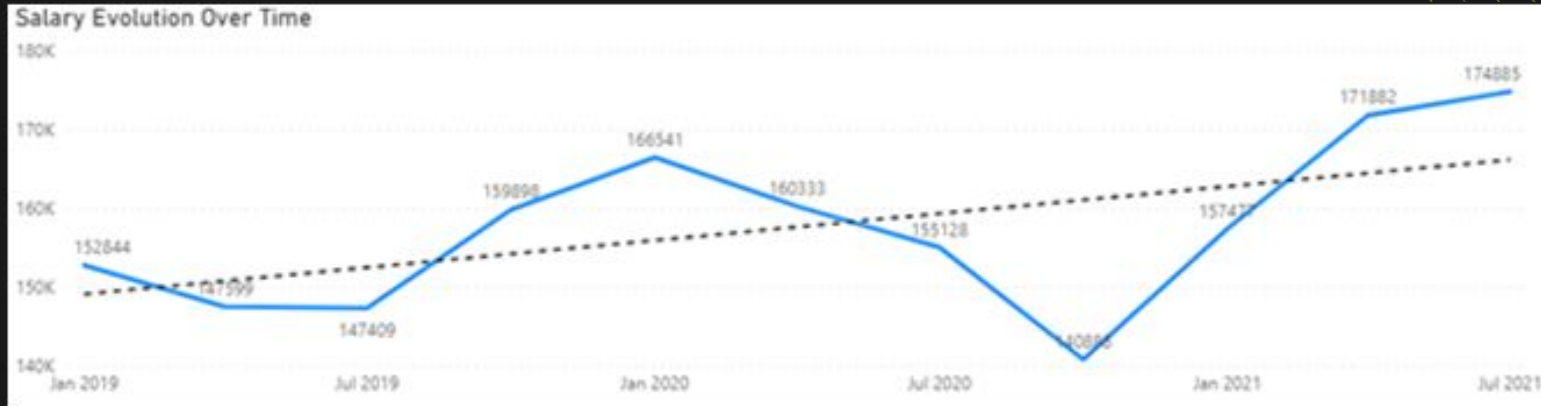


# Oportunidades na Área de Finanças Quantitativas





# Oportunidades na Área de Finanças Quantitativas





# Oportunidades na Área de Finanças Quantitativas

**Salaries and Bonuses in Quant Finance: broken down by role, seniority & region:**

<https://www.efinancialcareers.com/news/salaries-and-bonuses-in-quant-finance-broken-down-by-role-seniority-and-region>

**Position at Jane Street**

<https://www.janestreet.com/join-jane-street/position/4794171002/>



# O que é uma Estratégia de Investimento Quantitativa?

**Definição:** Estratégia que utiliza modelos matemáticos, estatísticos e algoritmos para tomar decisões de investimento.

**Base de Dados:** Depende de dados históricos e em tempo real para identificar padrões e sinais de mercado.

**Automatização:** As operações são executadas automaticamente, minimizando a intervenção humana e os vieses emocionais.

**Backtesting:** Os modelos são testados com dados passados para validar sua eficácia antes de serem aplicados em tempo real.

**Escalabilidade:** Permite a análise simultânea de grandes volumes de dados e múltiplos ativos.

**Aprimoramento Contínuo:** Pode incorporar técnicas de machine learning para refinar e adaptar os modelos conforme as condições de mercado mudam.



# Por onde começar a estratégia?

## Ativos

É necessário, antes de tudo, definir em que universo de ativos iremos trabalhar.

## Modelo

Aqui chegamos na etapa mais complexa da formação da estratégia, nesse ponto entram os modelos matemáticos e estatísticos que irão basear a nossa previsão.

## Dados

Talvez os dados sejam a etapa mais importante das estratégias quantitativas, afinal sem dados não há nenhum trabalho a ser feito.

## Estratégia

Sabendo o ativo que será trabalhado e o modelo adotado, é possível identificar características que corroboram para o melhor funcionamento de x estratégia.

## Backtesting

Depois de tudo definido, não temos como simplesmente chegar aplicando a estratégia, e aí entra o backtest, quando testaremos nosso modelo no passado, avaliando sua eficácia e acurácia.



# A Estratégia que Vamos Apresentar

## Estratégia

Compramos o primeiro percentil de ações com maior variação de preço e damos short, caso possível, nas ações do último percentil

## MSGARCH

A grande base da nossa estratégia é a previsão de volatilidade através do modelo econométrico MSGARCH

## Backtesting

Realizamos um backtest vetorizado sobre os dados adquiridos pós aplicação da estratégia

## LSTM

Modelo de IA aplicado nas ações filtradas por high ou low vol de acordo com o MSGARCH, utilizada para prever a variação de preços.

## Universo de ativos

S&P500







# Markov-Switching GARCH

O modelo MSGARCH combina dois conceitos importantes. Primeiro, ele usa um modelo de mudanças de regime (*Markov-Switching*), que permite que a série financeira alterne entre diferentes estados, como períodos de baixa e alta volatilidade. Depois, dentro de cada regime, ele aplica um modelo GARCH, que é uma forma de capturar como a volatilidade depende dos retornos passados. Isso torna o modelo mais flexível e realista, pois permite que a volatilidade siga padrões diferentes dependendo do regime em que o mercado se encontra.

O modelo MSGARCH **assume que os retornos  $r_t$  seguem um processo de volatilidade condicional que muda ao longo do tempo de acordo com um processo de Markov.** Ele pode ser escrito como:

$$r_t = \sigma_t \epsilon_t$$

onde:

$r_t$  é o retorno do ativo no tempo  $t$

$\sigma_t$  é a volatilidade condicional no tempo  $t$

$\epsilon_t$  representa choques aleatórios no mercado no tempo  $t$ , **assumindo que eles seguem uma distribuição específica (como normal ou t-student) para tornar o modelo tratável.**



# Markov-Switching GARCH

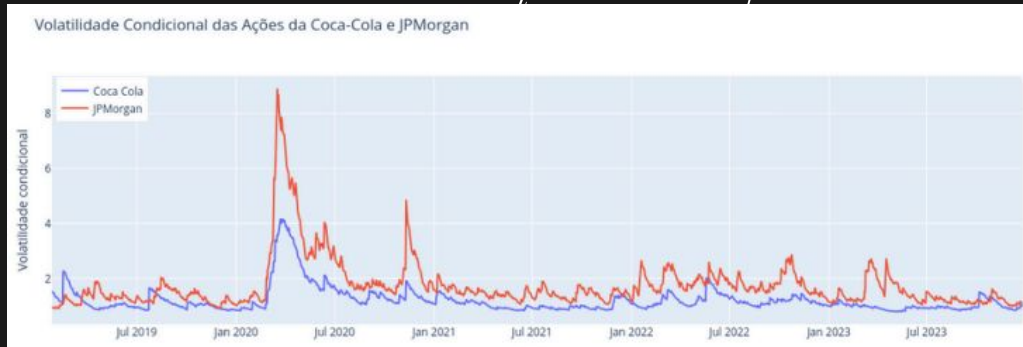
A grande diferença para um GARCH tradicional está na volatilidade  $\sigma_t$ , que agora depende de um regime oculto  $S_t$ :

$$S_t = \sigma_t^2 = \alpha S_t + \beta S_t r_{t-1}^2 + \gamma S_t \sigma_{t-1}^2$$

onde  $S_t$  é o estado do mercado no tempo  $t$ , determinado por uma cadeia de Markov de transição:

$$p_{ij} = P(S_t = j | S_{t-1} = i) = p_{ij}$$

onde  $p_{ij}$  representa a probabilidade de mudar do regime  $i$  para o regime  $j$ .





# Machine Learning aplicado a problemas de previsão

Modelos de **Machine Learning** são frequentemente usados em problemas de previsão devido à sua **capacidade de aprender padrões complexos a partir de dados históricos, sem a necessidade de especificar uma fórmula matemática explícita**. Esses modelos são particularmente úteis em problemas em que as relações entre variáveis são não lineares ou difíceis de modelar diretamente com abordagens tradicionais.

Modelos de **Machine Learning** são frequentemente aplicados para resolver dois grandes tipos de problemas de previsão:

1. **Classificação** – O modelo prevê categorias (ex.: prever se o mercado vai subir ou cair).
2. **Regressão** – O modelo prevê valores contínuos (ex.: prever o retorno esperado de um ativo).

Aqui, lidaremos com um problema de **regressão**, pois queremos prever o risco futuro de um ativo.



# Machine Learning aplicado a problemas de previsão

## O que é um modelo de Machine Learning?

Um modelo de **machine learning** é uma representação matemática que descreve a relação entre as variáveis de entrada (características ou "features") e as variáveis de saída (resultados ou "rótulos"). Ele é criado a partir de dados e treinado para identificar padrões ou regularidades que ajudam a fazer previsões ou decisões sobre dados novos. Como mencionamos anteriormente, não precisamos fornecer uma fórmula matemática específica para criar um modelo de machine learning, e isso é verdade porque **esses modelos têm a capacidade de aprender essas relações diretamente a partir dos dados**.

## Exemplo Intuitivo:

Imagine que você tem um conjunto de dados sobre imóveis. As entradas podem ser características como o tamanho do imóvel, a localização e o número de quartos, enquanto a saída seria o preço do imóvel. O modelo de **machine learning** tenta entender como essas características influenciam o preço, de forma que, quando você fornecer as características de um novo imóvel, o modelo consiga prever seu preço. Em outras palavras, o modelo aprende a criar uma função matemática que, ao receber como entrada as características do imóvel, gera como saída o preço correspondente.



# Machine Learning aplicado a problemas de previsão

## O que é uma Rede Neural e por que usá-las?

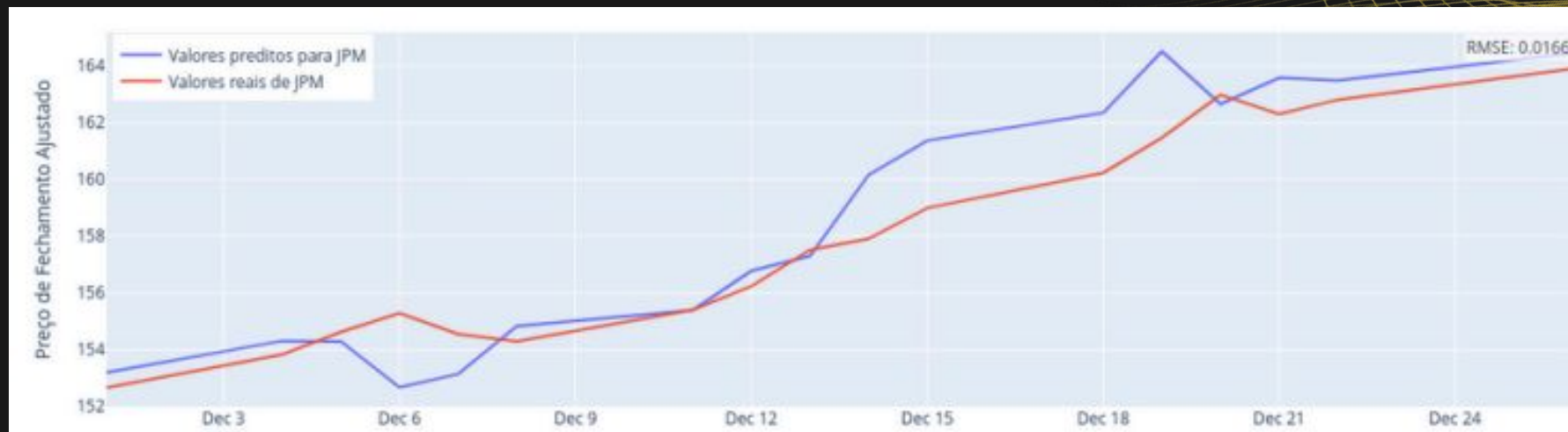
Redes neurais são modelos poderosos de **machine learning** inspirados no cérebro humano. Elas são particularmente eficazes para lidar com problemas complexos e dados **não lineares**, capturando relações difíceis de modelar com abordagens tradicionais. Essa capacidade de identificar padrões complexos é o que torna as redes neurais tão eficazes para prever fenômenos como o **preço de ativos**, que seguem dinâmicas altamente não lineares e dependem de múltiplos fatores interconectados.

Existem várias **arquiteturas** de redes neurais, cada uma com características específicas para diferentes tipos de tarefas. No contexto da **previsão de séries temporais**, as **redes neurais recorrentes (RNNs)** se destacam. Dentro desse grupo, a arquitetura **LSTM (Long Short-Term Memory)** é particularmente eficaz.

Redes LSTM são projetadas para capturar **dependências de longo prazo**, permitindo que o modelo compreenda padrões temporais mais complexos e faça previsões mais precisas, especialmente quando a **informação histórica** é crucial para prever o futuro.



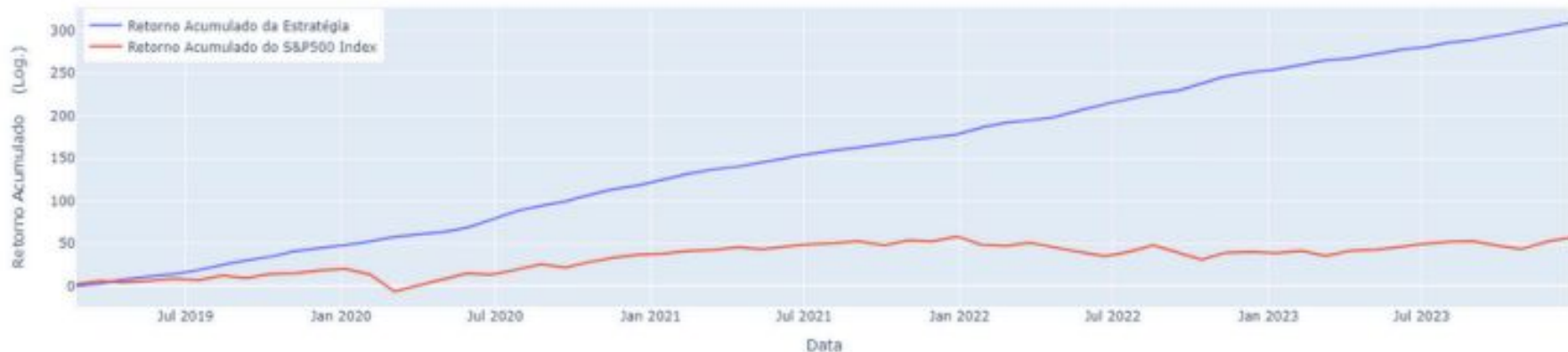
# Resultados





# Resultados

Retorno Acumulado da Estratégia vs. Retorno Acumulado do S&P500 Index





# Resultados

Retornos Mensais da Estratégia Vs. Retornos Mensais do S&P500 Index

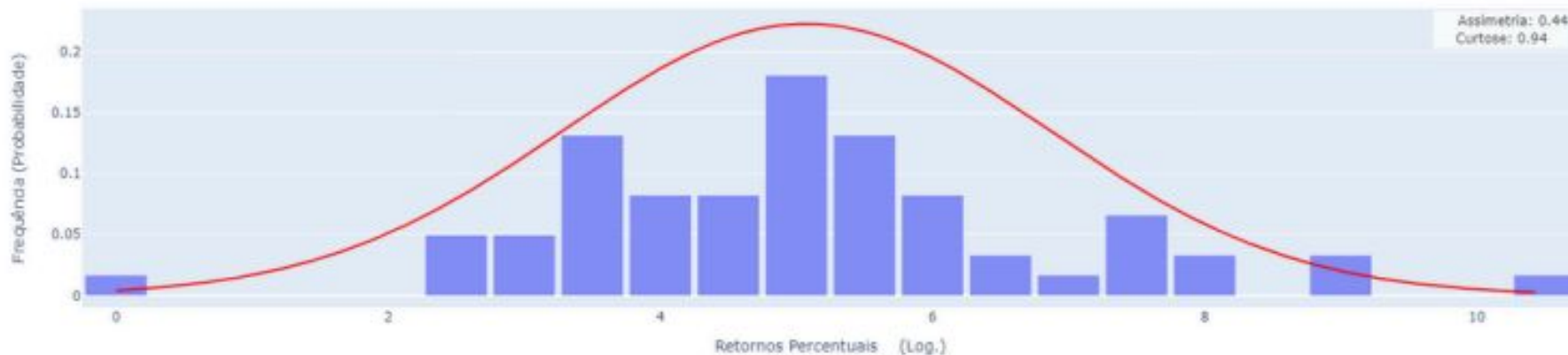






# Resultados

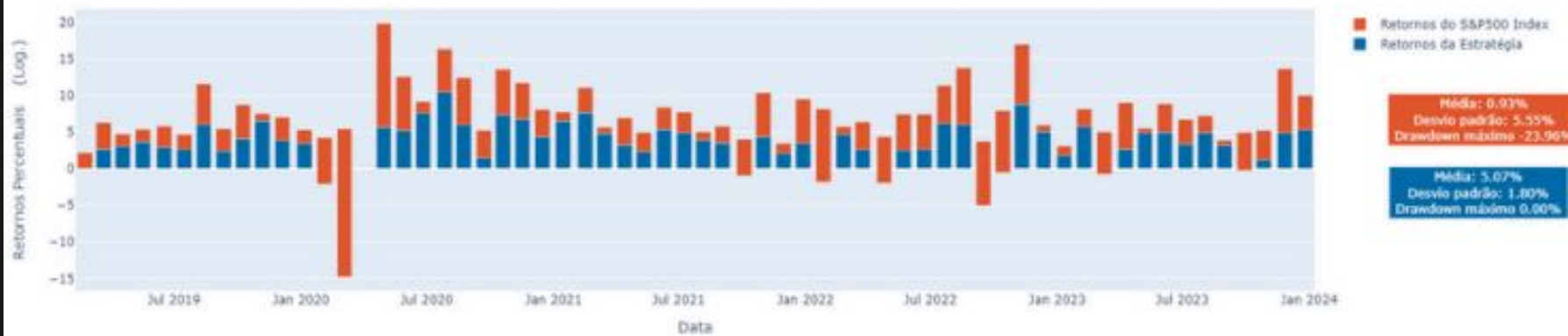
Histograma dos Retornos Mensais da Estratégia com Curva Normal Ajustada





# Resultados

Retornos Mensais da Estratégia vs. Retornos Mensais do S&P500 Index





# Resultados

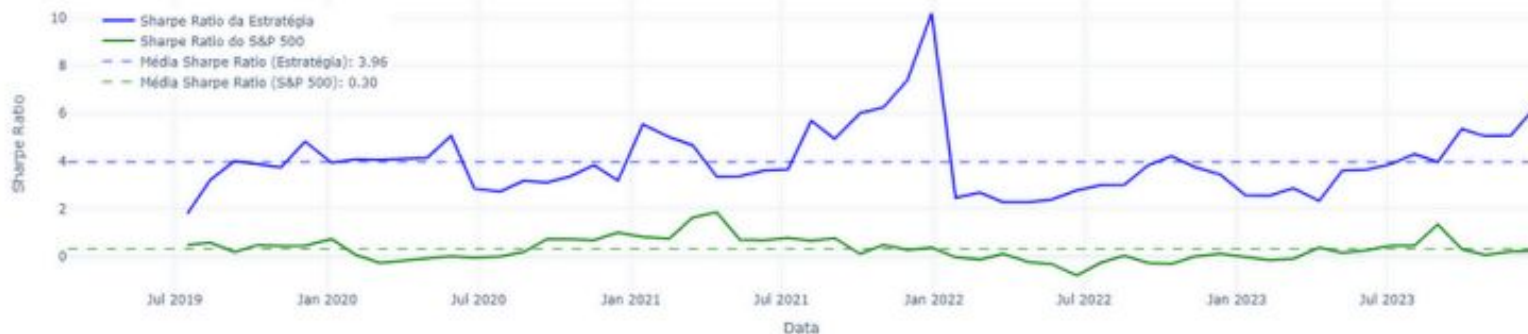
VaR e CVaR ao longo do tempo





# Resultados

Sharpe Ratio da Estratégia Vs. Sharpe Ratio do S&P 500 ao longo do tempo





## Resultados

A tabela ao lado apresenta um intervalo de erro da previsão, isto é, o quanto o retorno real divergiu do retorno previsto pela LSTM, os valores vieram das seis vezes que testamos a rede neural. Observando a tabela, podemos calcular que mais de 85% dos ativos tiveram seus preços previstos com um erro de no máximo 5%. Além disso, não há a presença de outliers com uma alta margem de erro, reforçando a robustez do nosso modelo.

Intervalo da Margem de Erro	Quantidade de Resultados
0 - 1%	37
1 - 2%	52
2 - 5%	124
5 - 10%	34
10 - 20%	3
20 - 50%	0



# Problemas de estratégias quantitativas

## Overfitting:

- Modelos que se ajustam demais aos dados históricos podem não generalizar bem para dados futuros.

## Qualidade dos Dados:

- Dados incompletos, imprecisos ou desatualizados podem comprometer a eficácia do modelo.

## Mudanças de Regime:

- Alterações nas condições do mercado podem tornar os modelos obsoletos rapidamente.

## Risco de Modelagem:

- Escolha inadequada de técnicas, parâmetros ou variáveis pode levar a previsões erradas.

## Custos de Transação e Slippage:

- Operações frequentes podem aumentar os custos e reduzir a rentabilidade real.

## Complexidade e Interpretação:

- Modelos complexos podem ser difíceis de interpretar e ajustar, dificultando a identificação de erros ou vieses.