

# Otimização de Carteira de Investimentos Utilizando Algoritmo Genético

**Autores:** Grupo 89

**Araguacy Pereira** - [araguacybp@yahoo.com.br](mailto:araguacybp@yahoo.com.br)

**Robson Calixto** - [robsoncalixto@gmail.com](mailto:robsoncalixto@gmail.com)

**Vinícius Costa** - [mcostavini98@gmail.com](mailto:mcostavini98@gmail.com)

**Instituição:** FIAP - Pós-Graduação IA para Devs

**Data:** 25 de Julho de 2025

---

## Sumário

1. [Introdução](#)
    - [1.1. Perguntas de Pesquisa](#)
    - [1.2. Contexto do Problema](#)
    - [1.3. Métodos de Otimização Financeira](#)
  2. [Metodologia](#)
    - [2.1. Coleta e Pré-processamento dos Dados](#)
    - [2.2. Definição do Portfólio e Métricas de Risco](#)
    - [2.3. O Algoritmo Genético para Otimização](#)
  3. [Testes e Avaliação](#)
    - [3.1. Configuração do Experimento](#)
    - [3.2. Análise de Convergência](#)
  4. [Análise dos Resultados](#)
    - [4.1. Composição do Portfólio Otimizado](#)
    - [4.2. Comparativo de Risco e Retorno](#)
  5. [Discussão](#)
  6. [Conclusão](#)
  7. [Trabalhos Futuros](#)
  8. [Referências Bibliográficas](#)
  9. [Anexos](#)
-

# 1. Introdução

## 1.1. Perguntas de Pesquisa

### O que está sendo otimizado?

O objetivo é maximizar o retorno ajustado ao risco de um portfólio de investimento. Isso implica em buscar a combinação ideal de aporte entre as ações selecionadas, de forma a controlar volatilidade. A participação (ou pesos) de cada ativo no portfólio é a variável que queremos otimizar.

### Qual é a representação da solução (genoma)?

A porcentagem de capital alocada entre os diferentes ativos disponíveis na carteira é o genoma, onde cada gene corresponde a proporção de capital em um ativo específico.

### Qual é a função de fitness?

A função considera a média ponderada de retorno dos ativos menos o CVaR ajustado por um fator de aversão ao risco. Quando adicionado um percentual de aversão ao risco, o CVaR é considerado para diversificar e mitigar o risco não sistemático.

Portanto, a função de fitness busca uma combinação de máximo retorno com o menor risco (volatilidade), sendo o risco medido pelo CVaR.

**VaR (Valor em Risco):** Indica a perda máxima esperada para um determinado nível de confiança e horizonte de tempo.

**CVaR (Valor Condicional em Risco):** Indica a perda máxima esperada para um determinado nível de confiança e horizonte de tempo, considerando especificamente os cenários de cauda (piores casos).

### Qual é o método de seleção?

O método de torneio foi escolhido como o ideal para o projeto por priorizar indivíduos com fitness mais alto, mantendo diversidade, mas favorecendo os melhores. Desta forma locais subótimos são evitados, permitindo que o algoritmo explore melhor o espaço de soluções até alcançar o objetivo.

### Qual método de crossover foi implementado?

O método implementado é o de corte único por posição, adaptado para um dicionário de pesos de ativos. Ele consiste em dividir os ativos ao meio e combinar partes dos pesos de dois portfólios diferentes para gerar dois novos filhos. A primeira metade dos ativos de um portfólio

é combinada com a segunda metade do outro, e vice-versa, criando variações a partir de indivíduos existentes.

## Qual é o método de inicialização?

Para iniciar a população, foi utilizada a distribuição aleatória de pesos normalizados para garantir que somem 1.

## Qual o critério de parada?

Número máximo de gerações definido ou ausência de melhoria significativa no fitness após um certo número de iterações definida por parâmetro.

## 1.2. Contexto do Problema

A alocação de capital em mercados financeiros é um dos problemas mais clássicos e desafiadores no campo das finanças quantitativas. Investidores buscam constantemente a construção de carteiras (ou portfólios) que não apenas maximizem os retornos, mas também gerenciem e minimizem os riscos associados. A Teoria Moderna do Portfólio, introduzida por Harry Markowitz, estabeleceu as bases para a diversificação como uma ferramenta para otimizar o trade-off entre risco e retorno.

No entanto, encontrar a alocação ótima de ativos em um portfólio é um problema de otimização complexo, especialmente quando se lida com um grande número de ativos e restrições não lineares. Métodos tradicionais podem ser computacionalmente intensivos ou ficar presos em ótimos locais.

Este projeto aborda o desafio de otimizar uma carteira de investimentos composta por ações da bolsa de valores brasileira (B3). O objetivo é utilizar **Algoritmo Genético (AG)**, uma meta-heurística inspirada na teoria da evolução de Darwin, para encontrar a distribuição de capital entre os ativos que otimiza uma função objetivo baseada em retorno e risco. A variável a ser otimizada é o vetor de pesos que representa a porcentagem de capital alocada em cada ativo do portfólio.

## 1.3. Métodos de Otimização Financeira

A otimização de portfólios pode ser abordada por diversas técnicas, desde a programação quadrática usada no modelo de Markowitz até métodos mais sofisticados. As principais métricas de risco utilizadas para avaliar a qualidade de um portfólio incluem:

- **Volatilidade (Desvio Padrão):** A medida clássica de risco, que quantifica a dispersão dos retornos de um ativo.

- **Valor em Risco (VaR):** Estima a perda máxima esperada de um portfólio em um determinado horizonte de tempo e com um certo nível de confiança. Apesar de popular, o VaR não é uma medida de risco coerente, pois falha na propriedade da subaditividade, o que significa que a diversificação pode, paradoxalmente, aumentar o VaR.
- **Valor em Risco Condicional (CVaR):** Também conhecido como *Expected Shortfall*, o CVaR mede a perda média esperada, dado que a perda excede o VaR. O CVaR é uma medida de risco coerente e mais conservadora, pois foca na magnitude das perdas nos piores cenários.

Neste projeto, o **CVaR** foi escolhido como a principal métrica de risco a ser minimizada, devido à sua robustez teórica e sua capacidade de capturar o risco de cauda, que é crucial em mercados financeiros. Os Algoritmos Genéticos são particularmente adequados para esse tipo de problema, pois podem explorar um vasto espaço de soluções de forma eficiente, sem fazer suposições sobre a convexidade da função objetivo, lidando bem com a complexidade do CVaR.

## 2. Metodologia

### 2.1. Coleta e Pré-processamento dos Dados

Os dados utilizados neste estudo consistem em cotações históricas diárias de fechamento de um conjunto de ações listadas na B3.

- **Fonte de Dados:** Os dados foram obtidos através da biblioteca `yfinance` do Python, que acessa a API pública do Yahoo Finance.
- **Seleção de Ativos:** Para focar o processo de otimização em ativos com bom desempenho histórico, foi realizado um filtro inicial. O sistema seleciona as 20 ações com o maior retorno médio nos últimos dois anos a partir de uma lista pré-definida de empresas brasileiras.
- **Pré-processamento:** A partir dos preços de fechamento, foram calculados os retornos diários percentuais, que servem como entrada para o cálculo das métricas de retorno e risco do portfólio.

### 2.2. Definição do Portfólio e Métricas de Risco

Um portfólio é definido por um conjunto de ativos e os respectivos pesos que indicam a proporção do capital total investido em cada um. A principal restrição é que a soma de todos os pesos deve ser igual a 1 (ou 100%).

A qualidade de um portfólio candidato é avaliada por uma **função de fitness**, que foi projetada para equilibrar o retorno esperado e o risco (CVaR):

$$\text{Fitness} = (1 - \text{Fator de Aversão ao Risco}) * \text{Retorno Esperado} - \text{Fator de Aversão ao Risco} * \text{CVaR}$$

- **Retorno Esperado:** É a média ponderada dos retornos médios históricos de cada ativo no portfólio.
- **CVaR:** Calculado com um nível de confiança de 95%, representa a perda média esperada nos 5% piores dias de negociação.
- **Fator de Aversão ao Risco:** Um parâmetro (entre 0 e 1) que permite ao investidor definir a importância do risco na otimização. Um valor maior indica que o investidor prefere minimizar o risco, mesmo que isso signifique um retorno menor.

## 2.3. O Algoritmo Genético para Otimização

Para encontrar o portfólio ótimo, foi implementado um Algoritmo Genético com os seguintes componentes:

- **Representação (Genoma):** Cada indivíduo (cromossomo) na população representa uma solução candidata, ou seja, um portfólio. O genoma é um dicionário onde as chaves são os *tickers* dos ativos e os valores são seus respectivos pesos. Uma função de normalização garante que a soma dos pesos seja sempre 1.
- **População Inicial:** A primeira geração de portfólios é criada aleatoriamente, atribuindo pesos aleatórios a cada ativo e normalizando-os.
- **Seleção (Torneio):** Para selecionar os pais da próxima geração, utiliza-se o método de **seleção por torneio**. Três indivíduos são escolhidos aleatoriamente da população, e os dois com o maior valor de fitness são selecionados como pais. Este método equilibra a pressão seletiva com a diversidade da população.
- **Crossover (Ponto de Corte Único Adaptado):** Os pais selecionados geram filhos através de um operador de crossover. O conjunto de ativos é dividido ao meio, e um filho é criado combinando a primeira metade dos pesos de um pai com a segunda metade do outro. Os pesos do filho resultante são normalizados.
- **Mutação (Perturbação Gaussiana):** Para introduzir novidade genética e evitar a convergência prematura, um operador de mutação é aplicado. Com uma certa probabilidade, o peso de um gene (ativo) é ligeiramente modificado por um valor aleatório.
- **Elitismo:** Para garantir que as melhores soluções encontradas não sejam perdidas, uma estratégia de elitismo é adotada. Os 10% melhores indivíduos de cada geração são transferidos diretamente para a próxima, sem sofrer crossover ou mutação.
- **Critério de Parada:** O algoritmo executa por um número pré-definido de gerações (ex: 50).

## 3. Testes e Avaliação

### 3.1. Configuração do Experimento

O algoritmo foi executado utilizando uma interface web interativa construída com a biblioteca **Streamlit**. Isso permitiu a configuração dinâmica dos hiperparâmetros do AG e a visualização em tempo real dos resultados.

Os principais parâmetros utilizados foram:

- **Tamanho da População:** 10 indivíduos
- **Número de Gerações:** 50
- **Taxa de Mutação:** 20%
- **Taxa de Crossover:** 50%
- **Fator de Aversão ao Risco:** 0.5 (equilíbrio entre risco e retorno)
- **Percentual de Elite:** 10%

### 3.2. Análise de Convergência

A eficácia do algoritmo foi monitorada através de um gráfico de convergência, que plota o fitness do melhor indivíduo e o fitness médio da população ao longo das gerações.

Observou-se que o algoritmo converge rapidamente nas primeiras gerações, com o fitness do melhor indivíduo aumentando significativamente. Após cerca de 20-30 gerações, o progresso tende a se estabilizar, indicando que a solução está próxima de um ótimo local ou global. A estratégia de elitismo se mostrou eficaz em garantir um progresso monotônico do melhor fitness.

## 4. Análise dos Resultados

### 4.1. Composição do Portfólio Otimizado

Ao final da execução, o algoritmo retorna o melhor portfólio encontrado. A composição deste portfólio representa a alocação de capital ótima segundo os critérios definidos. Tipicamente, o AG tende a concentrar os investimentos em um número limitado de ativos que apresentam a melhor combinação de alto retorno e baixo CVaR, validando o princípio da diversificação seletiva.

A interface web apresenta um gráfico de pizza com a distribuição percentual dos ativos no portfólio final, facilitando a interpretação do resultado pelo usuário.

### 4.2. Comparativo de Risco e Retorno

O resultado final inclui as métricas de desempenho do portfólio otimizado:

- **Fitness Final:** O valor máximo da função objetivo alcançado.

- **Retorno Esperado Anualizado:** O retorno projetado do portfólio.
  - **CVaR Anualizado:** O risco de cauda do portfólio.
- Esses valores permitem ao investidor avaliar se o trade-off entre o risco assumido e o retorno esperado está de acordo com seu perfil.

## 5. Discussão

A utilização de Algoritmos Genéticos para a otimização de portfólios demonstrou ser uma abordagem robusta e flexível. Diferente de métodos clássicos que podem exigir simplificações no modelo de risco, os AGs operam diretamente sobre a função de fitness, mesmo que ela seja complexa e não-convexa, como é o caso da otimização baseada em CVaR.

A capacidade de explorar um vasto espaço de soluções permite que o algoritmo escape de ótimos locais e encontre soluções de alta qualidade que seriam difíceis de alcançar com otimizadores tradicionais. A natureza estocástica do AG também reflete a incerteza inerente aos mercados financeiros.

No entanto, a qualidade da solução depende da calibração dos hiperparâmetros do AG (tamanho da população, taxas de mutação/crossover). Uma configuração inadequada pode levar à convergência prematura ou a uma exploração ineficiente do espaço de busca. A implementação de taxas de mutação e crossover adaptativas poderia ser uma melhoria futura para refinar ainda mais o processo.

## 6. Conclusão

Este projeto demonstrou com sucesso a aplicação de Algoritmos Genéticos para resolver o problema complexo de otimização de carteiras de investimentos. A implementação foi capaz de identificar portfólios que otimizam o retorno ajustado ao risco, utilizando o CVaR como uma medida de risco amplamente utilizada no mercado financeiro.

A abordagem se mostrou viável e produziu resultados coerentes e interpretáveis, fornecendo uma ferramenta poderosa para a tomada de decisão em investimentos. A combinação de uma meta-heurística poderosa com uma interface interativa (Streamlit) cria uma solução prática e acessível para a análise financeira.

## 7. Trabalhos Futuros

Para a evolução deste projeto, recomenda-se:

- **Backtesting Robusto:** Implementar um sistema de backtesting que avalie o desempenho do portfólio otimizado em dados "fora da amostra" (períodos futuros), para validar a eficácia preditiva da estratégia.

- **Otimização Multi-Objetivo:** Evoluir o AG para um algoritmo de otimização multi-objetivo (como o NSGA-II) para gerar uma **Frenteira Eficiente** completa, mostrando o portfólio ótimo para cada nível de risco, em vez de um único ponto.
  - **Parâmetros Adaptativos:** Implementar mecanismos que ajustem dinamicamente as taxas de mutação e crossover durante a execução, melhorando a capacidade de exploração e exploração do algoritmo.
  - **Inclusão de Mais Restrições:** Adicionar outras restrições do mundo real, como limites de alocação por setor, liquidez mínima dos ativos e custos de transação.
- 

## 8. Referências Bibliográficas

1. \*Documento de Referência: [Otimização de Redes Neurais utilizando Algoritmos Genéticos para Previsão de Sobrevivência no Titanic](#). FIAP, 2024.
  2. VANELI, Daniel Mognato. [OTIMIZAÇÃO DE PORTFÓLIO COM ATIVOS DO IBOVESPA USANDO ALGORITMOS GENÉTICOS](#). Trabalho de Conclusão de Curso – Ifes: Instituto Federal do Espírito Santo, Espírito Santo, 2024.
  3. Maxwell. Modelagem estocástica e medida de risco. [https://www.maxwell.vrac.puc-rio.br/26820/26820\\_6.PDF](https://www.maxwell.vrac.puc-rio.br/26820/26820_6.PDF), Tese de Doutorado - PUC RIO, Rio de Janeiro. Acesso: 14 de Julho. 2025
  4. BORGES, Bruno. *Otimizando carteiras de investimentos com Data Science*. Medium, 2021. Disponível em: <https://medium.com/ensina-ai/otimizando-carteiras-de-investimentos-com-data-science-f545dbe30bae>.
  5. REIS, Tiago. *Teoria de Markowitz: como calcular a relação de risco e retorno*. Suno, 2021. Disponível em: <https://www.suno.com.br/artigos/teoria-de-markowitz/>
  6. Empresas Listadas B3. Dados abertos da B3. Disponível em: [https://www.b3.com.br/pt\\_br/produtos-e-servicos/negociacao/renda-variavel/empresas-listadas.htm](https://www.b3.com.br/pt_br/produtos-e-servicos/negociacao/renda-variavel/empresas-listadas.htm) Acesso: 08 de julho. 2025.
- 

## 9. Anexos

- **Repositório do Código-Fonte:** [https://github.com/robsoncalixto/wallet-optimization\\_ag](https://github.com/robsoncalixto/wallet-optimization_ag)
- **Aplicação Web (Streamlit):** <https://wallet-optimizationag-cm6gbhspo4cjx3czsvj36u.streamlit.app/>