**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO SUL DE MINAS GERAIS - CAMPUS POÇOS DE CALDAS**

**CURSO ENGENHARIA DE COMPUTAÇÃO**

**ELIEZER MARQUES MAFRA**

**GESTÃO DE PORTFÓLIO DE CRIPTOMOEDAS MULTINÍVEL E MULTIOBJETIVO UTILIZANDO ALGORITMO GENÉTICO  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
POÇOS DE CALDAS**

**2022**

**ELIEZER MARQUES MAFRA  
  
  
  
  
  
  
  
  
  
  
  
GESTÃO DE PORTFÓLIO DE CRIPTOMOEDAS UTILIZANDO ALGORITMO GENÉTICO MULTINÍVEL**

Trabalho de Conclusão de Curso apresentado ao Curso Engenharia de Computação do Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais - Campus Poços de Caldas, como requisito parcial à obtenção do título de Bacharel em Engenharia de Computação.

Orientador: Prof. Dr. Humberto Vargas Duque

**POÇOS DE CALDAS**

**2022**

# INTRODUÇÃO

## PROBLEMATIZAÇÃO

Bancos e instituições financeiras ao redor de todo mundo têm aceitado criptomoedas como meio de pagamento ou como meio de investimento com índices de retorno elevados. Os preços das criptomoedas são muito mais voláteis quando comparados com os tradicionais ativos de investimento, apresentando volatilidades diárias em seu preço de até 1000% em alguns casos de valorização. Isso torna o investimento em criptomoedas atrativo ao mesmo tempo que expõe o investidor a um risco significativo (BOIKO et al., 2021). O investimento neste tipo de ativo pode ser monitorado e gerenciado de forma a se diminuir os seus riscos visando melhorar o lucro do portfólio?

## OBJETIVOS

### Objetivo Geral

Implementar um algoritmo que gerencie riscos de um portfólio de investimentos em criptomoedas.

### Objetivo Específico

* + - Desenvolver um algoritmo genético multiobjetivo com função de diminuir o risco e aumentar o retorno.
    - Gerar relatórios de investimentos com a sugestão de alocação em cada criptoativo.
    - Implementar algoritmo genético que dê as melhores configurações paramétricas para se gerar portifólios com risco reduzido com retornos consideráveis.

## HIPÓTESE

É possível estabelecer um portfólio de investimento em criptomoedas com um índice de risco moderado com um retorno satisfatório.

# JUSTIFICATIVA

## 2.1 CONTEXTO HISTÓRICO (BITCOIN)

Bitcoin foi a primeira criptomoeda do mundo, lançada em 2009, com o objetivo de ser uma troca de valores global sem intermediários, e propondo uma nova forma de se fazer finanças sem interferências de bancos ou governos, por isso ele é chamado de descentralizado. Ele é um meio seguro por ser desenvolvido a partir da complexa criptografia (por isso é chamado de cripto), e ter seus registros validados e registrados em um livro razão virtual, chamado de Blockchain. Por meio de diversos participantes autônomos da sua rede as negociações são validadas sem a presença de uma autoridade central, como dito antes.

A maneira mais popular e segura para se investir em BTC, de qualquer lugar, é através de corretoras de criptomoedas, chamadas de exchanges. Essas empresas financeiras disponibilizam plataformas virtuais, em computadores ou smartphones, que permitem você comprar e negociar Bitcoin e outros ativos digitais. Essas plataformas não precisam ser nacionais. Um bom exemplo de exchange é a Binance e a Bitmex, que foram as exchanges de onde os dados desse trabalho foram obtidos. (COMO INVESTIR..., 2021)

## 2.2 MOTIVAÇÃO

Em meio aos efeitos da pandemia do coronavírus, a taxa média de juros bancários reduziu em todo o mundo no ano de 2020. (DE CASTRO, 2021). O Bitcoin, considerado a principal criptomoeda do mercado teve uma valorização de 72% no começo do ano de 2021. Essa alta rentabilidade atraiu muitos investidores, principalmente pessoas físicas, que buscavam aplicações mais rentáveis visto a taxa de juros tão baixa da época. Com a Selic em 2%, valor da taxa em fevereiro de 2021, a rentabilidade de ativos mais tradicionais, de renda fixa, ficou menos atrativa. (BOURGUIGNON, 2021)

No entanto, devido à alta volatilidade das criptomoedas muitos investidores se mostram céticos a este tipo de investimento devido ao seu alto risco. A partir de análises feitas no site TradingView®, o Bitcoin teve uma valorização de 1020% de agosto de 2020 à agosto de 2021, para meios de comparação, o índice Ibovespa (IBOV) teve uma valorização de 13,62% no mesmo período. Um estudo realizado pela consultoria Grimpa revela que há mudanças ocorrendo nas carteiras dos investidores, muitos deles estão incluindo criptomoedas em seus ativos como forma de diversificar o portfólio e buscar robustez nos retornos. A pesquisa, que ouviu 500 homens e mulheres com mais de 18 anos ao redor do País, mostra que 96% dos entrevistados já ouviram falar em criptomoedas e que 41% deste grupo já investiu em ativos digitais em algum momento. (ESTADÃO E-INVESTIDOR, 2021)

Como a principal desvantagem das criptomoedas é a sua volatilidade e consequentemente o seu risco, esse trabalho visa expor um gerenciamento de portfólio que diminui esses índices de risco ao mesmo tempo que tenta buscar um retorno razoável através de análise e processamento de dados.

# METODOLOGIA

## COLETA DE DADOS

Inicialmente a coleta de dados se deu por meio de scripts (NISTRUP, 2019) escritos na linguagem Python, foram coletados dados de cotação de 258 criptomoedas, nas quais citam-se: Bitcoin (BTC), Ethereum (ETH), Cardano (ADA), Binance Coin (BNB), Riple (XRP), Dogecoin (DOGE), Polkadot (DOT), Solana (SOL) e etc. Essas cotações foram obtidas em uma granularidade de 1 dia. Nesse caso foi utilizado uma API disponibilizada pela Binance juntamente com pacotes de processamento e manipulação de arquivos para exportar esses dados em arquivos com extensão csv.

## SELEÇÃO DE DADOS

Sendo o Bitcoin a criptomoeda mais antiga e também a que possui maior liquidez, ela foi escolhida como base de tempo para nosso estudo. Então a base de dados começa em 17 de agosto de 2017 e decorre ao longo de 196 semanas. De todos os ativos, foram selecionados todos que, ao começarem a ter cotações, não pararam até o fim do período, isso foi feito para se evitar criptomoedas que não se consolidaram e pararam de ser negociadas. Não é viável se analisar essas criptomoedas pois, mesmo que elas estejam presentes no período de validação, não será possível se construir um portifólio com elas no período de testes, pois elas não estarão mais disponíveis para investimento.

## VALIDAÇÃO E TESTE DO MODELO

O Teste de Validação de Modelos é o procedimento de avaliação do bem-estar do desempenho dos modelos em relação aos dados reais. É essencial que o modelo seja validado considerando os aspectos e os componentes antes de introduzi-los no ecossistema de produção.

Modelos de aprendizado de máquina, modelos de deep learning e outros modelos de ciência de dados são como assistentes para gerar a saída correta se a pergunta exata for feita a esses assistentes. Mas como a confiabilidade proporcionada a esses modelos. É aí que entra o teste de validação de modelo. Ele fornece resultados confiáveis ​​gerados por esses modelos por uma comparação matemática e lógica com a saída real. O teste de validação de modelo é executado no limite do ajuste do modelo para experimentar diferentes dados de teste e treinamento e verifica a validade do modelo de maneira em loop. O processo de validade automatizado, é uma grande vantagem para verificar a confiabilidade do modelo de Inteligência Artificial. (SINGH GILL, 2018)

Os dados foram divididos em conjuntos de validação e teste, o período de validação foi um período de 20 semanas que precediam o período de teste, que compreendia as próximas 96 semanas. Com a adoção dessa abordagem, foi somente selecionados as melhores configurações paramétricas para o período de validação, e estas foram aplicadas ao período de teste.

## TEORIA MODERNA DE PORTFÓLIO E TEOREMA DA VARIÂNCIA MÍNIMA GLOBAL

Markowitz (1952, 1959) é o pai da teoria moderna de portfólio. Seu livro e artigo originais sobre o assunto delinearam claramente, pela primeira vez, a teoria moderna do portfólio. O livro é repleto de insights e sugestões que

antecipou muitos dos desenvolvimentos subsequentes no campo. Markowitz formulou o problema do portfólio como uma escolha da média e variância de um portfólio de ativos. Ele provou o teorema fundamental da carteira de variância média teoria, ou seja, manter a variância constante, maximizar o retorno esperado e manter o retorno esperado constante, minimizar a variância. Esses dois princípios levaram à formulação de uma fronteira eficiente a partir da qual o investidor poderia escolher sua carteira preferida, dependendo das preferências individuais de retorno de risco. A mensagem importante da teoria era que os ativos não podiam ser selecionados apenas em características que eram únicas para a segurança. Em vez disso, um investidor tinha que considerar como cada título co-movimentado com todos os outros títulos. Além do mais, levar esses co-movimentos em consideração resultou na capacidade de construir um portifolio que tinha o mesmo retorno esperado e menor risco do que uma carteira construída ignorando as interações entre os títulos.

Considerando apenas o retorno médio e a variância do retorno de uma carteira é, de fato, uma simplificação relativa à inclusão de momentos adicionais que podem descrever de forma mais completa a distribuição dos retornos da carteira.

No entanto, a teoria da variância média permaneceu como o pilar da teoria moderna de portfólio apesar dessas alternativas. Essa persistência não é devido ao realismo das premissas de distribuição de utilidade ou retorno que são necessárias para que seja correta. Em vez disso, acredita-se que há duas razões para sua persistência. Em primeiro lugar, a própria teoria da variância média coloca grandes requisitos de dados sobre o investidor, e não há evidências de que adicionar momentos adicionais melhore a conveniência do portfólio selecionado. Em segundo lugar, as implicações da teoria do portfólio de variância média são bem desenvolvidas, amplamente conhecidas e têm grande apelo intuitivo. Profissionais que nunca executaram um otimizador aprenderam que correlações, bem como médias e variâncias são necessárias para entender o impacto de adicionar um título a um portfólio. (ELTON, et al, 1997).

Para cada criptomoeda, o risco pode ser calculado pela variância da cotação durante o período especificado (FU, et al, 2021).

|  |  |
| --- | --- |
|  | (1) |

Onde é o número de dias no período, é a cotação da criptomoeda no dia , e é a média das cotações no período .

Ao se analisar a covariância, ela demonstra a similaridade da flutuação do preço das criptomoedas. Por exemplo, a covariância entre duas criptomoedas A e B:

|  |  |
| --- | --- |
|  | (2) |

Onde é o número de dias no período, é a cotação da moeda no dia , é a cotação da moeda no dia , é a média das cotações da moeda e da moeda no período .

Entretanto, como um portifólio consiste de um número de ativos, no caso de criptomoedas, o risco do portifólio será calculado como:

|  |  |
| --- | --- |
|  | (2) |

Onde e é a peso das criptomoedas e no portifólio, é a variância da moeda e é a covariância entre a moeda e (FU, et al, 2021). Essa covariância representa o risco de se colocar os ativos e no portifólio. (MARKOWITZ, 1952, 1959)

Assim, para se obter o portifólio de menor risco, deve-se minimizar o risco do portifólio com:

|  |  |
| --- | --- |
|  | (3) |

Sujeito a seguinte restrição:

|  |  |
| --- | --- |
|  | (4) |

Pois, o somatório de todas as proporções de cada ativo devem obrigatoriamente somar , pois nesse trabalho, não se considera operações a descoberto.

Não é interessante obtermos simplesmente o portifólio com o menor mínimo global, pois o objetivo desse gerenciamento de portifólio não é simplesmente obter a carteira com o menor risco, mas sim, contrabalancear o risco com o retorno.

Aqui, vê-se a necessidade de implementação de um algoritmo que encontre não o mínimo global, conforme é possível encontrar com algoritmos exatos, mas mínimos locais, que ofereçam retornos mais atrativos. Esses retornos locais são possíveis de serem encontrados utilizando um algoritmo estocástico, tal como um Algoritmo Genético.

## ALGORITMOS GENÉTICOS

Algoritmos Genéticos (AGs) podem ser definidos como procedimentos de pesquisa baseados na genética e na seleção natural de espécies. Assim como no ambiente natural, em um AG existe um grupo de candidatos a soluções, também conhecidos como indivíduos, que competem entre si para garantir sua própria sobrevivência e garantir que suas características sejam passadas de geração em geração. (GOLDBERG; HOLLAND, 1988)

O algoritmo possui 4 etapas:

* + - * 1. Fitness: é usado para determinar a qualidade de um indivíduo.
        2. Seleção: é a etapa em que a população é avaliada por fitness e ordenada.
        3. Crossover: é onde ocorre o crossover, e esta etapa é executada fazendo com que os indivíduos como pais criem uma prole seguindo um algoritmo específico.
        4. Mutação: tem a função de modificar as características individuais por uma perturbação aleatória.

O algoritmo funciona com a estrutura da *Figura 1*, onde uma população pode ser definida como um conjunto de cromossomos, onde cada cromossomo é composto por genes. O tamanho dessa população é parte da configuração paramétrica do problema enquanto o tamanho do cromossomo é o peso percentual para cada ativo presente no portifólio. Os cromossomos podem ser chamados também de indivíduos, pois são indivíduos de uma população. A população inicial é definida aleatoriamente seguindo o critério imposto pela equação (4).



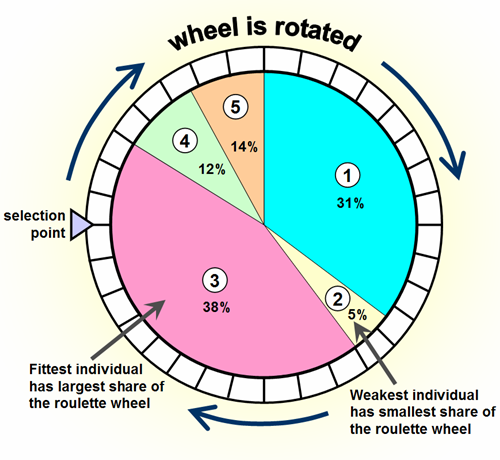
Figura 1: Estrutura Algoritmo Genético  
Fonte: (MALLAWAARACHCHI, 2017)

Para realização do crossover, primeiramente se seleciona os cromossomos que serão os pais para gerar um novo filho para a próxima geração. O método de seleção escolhido para uma melhor solução do problema foi a roleta. A roleta estabelece uma probabilidade de ser selecionado de acordo com a fitness do indivíduo. A equação *(1)* define a probabilidade de cada cromossomo ser selecionado.

|  |  |
| --- | --- |
|  | (1) |

Onde é a probabilidade de cada cromossomo ser igual à frequência cromossômica dividida pela soma de todos os fitness. Pode-se fazer uma analogia considerando o algoritmo de seleção da roleta é como um gráfico de pizza. Cada indivíduo tem um valor de aptidão e a soma de todos é o círculo. Portanto, a probabilidade de selecionar um parceiro em potencial depende de sua aptidão em relação ao resto. A *Figura 2* a seguir mostra que a probabilidade de selecionar cada um deles depende de quanto espaço eles ocupam na torta. (ROULETTE..., 2019)

**Roleta é rotacionada**



**Indivíduo com menor fitness tem uma menor porção da roleta**

**Indivíduo com maior fitness tem uma maior porção da roleta**

**Ponto de seleção**

Figura 2: Ilustração do algoritmo da roleta  
Fonte: (ROULETTE..., 2019)

Vale ressaltar também que foi implementado elitismo no crossover das populações, onde indivíduos que estivessem dentro elite, sempre são passados para a próxima geração e não sofrem mutação. Nesse caso é definido uma taxa de elitismo, que contém a taxa percentual da população com melhor fitness que é passada para a próxima geração.

A mutação é o processo mais simples, onde é selecionado um indivíduo ao acaso e um gene ao acaso e o substitui por um valor aleatório, novamente, é sempre aplicado o critério da equação (4).

Dentro do algoritmo genético, foi definido o conjunto de variáveis que formam a configuração paramétrica que é composta por:

* Tamanho da população: define o número de indivíduos de todas as gerações daquela execução do algoritmo.
* Probabilidade de mutação do gene: define a probabilidade de mutação de um gene.
* Probabilidade de mutação do indivíduo: define a probabilidade de mutação de um indivíduo.
* Probabilidade de crossover: define a probabilidade de cada indivíduo sofrer crossover.
* Taxa de elitismo: define a percentagem de indivíduos da população a serem mantidos na elite, sem sofrer crossover e mutação.
* Número de gerações: define o critério de parada pelo número de gerações da população.
  1. **ALGORITMO MULTINÍVEL**

O paradigma multinível refere-se ao processo de dividir problemas grandes e complexos em menores, que são muito mais fáceis de resolver, e então trabalhar para trás em direção à solução do problema original, usando a solução alcançada a partir de um nível filho como solução inicial para o nível dos pais. (BOUHMALA, 2017)

Nesse trabalho, foi implementado em um primeiro nível uma análise combinatória de configurações paramétricas para a minimização do risco como função objetivo do algoritmo genético.

Dos resultados desse algoritmo genético para o período de validação, foi implementado, em um segundo nível, a seleção, dentre essas configurações paramétricas, os 50 maiores retornos futuros considerando um período posterior a semana de teste de 7 dias e as que obtiveram os 50 menores desvios padrões.

O período de validação se iniciou na semana 80 da base de dados, pois em um período anterior a isso, havia muito poucos criptoativos para se montar um portifólio. A semana 80 foi selecionada empiricamente, por ter aproximadamente 50 ativos, que poderiam compor a carteira de investimentos. O período de validação se estendeu até a semana 99. O período de testes se iniciou na semana 100 e decorreu até a semana 196.

# ANÁLISE DOS RESULTADOS

//INSERIR RESULTADOS

# CONSIDERAÇÕES FINAIS

# REFERÊNCIAS

MARKOWITZ, Harry M. Portfolio Selection. **The Journal of Finance**, New York, v. 7, n. 1, p. 77-91, 1 mar. 1952. Disponível em: https://www.math.hkust.edu.hk/~maykwok/courses/ma362/07F/markowitz\_JF.pdf. Acesso em: 3 fev. 2022.

MARKOWITZ, H. Portfolio Selection: Efficient Diversification of Investment. **Cowles Foundation Monograph**, n. 16. New York: John Wiley E Sons, Inc, 1959. Disponível em: https://cowles.yale.edu/sites/default/files/files/pub/mon/m16-all.pdf. Acesso em: 3 fev. 2022.

BLOOMBERG. **Bitcoin é colocado em categoria de máximo risco sob plano de reguladores.** [S. l.]: Bloomberg, 2021. Disponível em: https://www.infomoney.com.br/mercados/bitcoin-e-colocado-em-categoria-de-maximo-risco-sob-plano-de-reguladores/. Acesso em: 29 jun. 2021.

GOLDBERG, David E.; HOLLAND, John H. Genetic Algorithms and Machine Learning. **Kluwer Academic Publishers**, [s. l.], 1988. Disponível em: https://link.springer.com/article/10.1023/A:1022602019183. Acesso em: 29 jun. 2021.

HOLLAND, John H. Genetic Algorithms. **Scientific American**, [s. l.], v. 267, ed. 1, p. 66-73, jul 1992. Disponível em: https://www.jstor.org/stable/24939139. Acesso em: 29 jun. 2021.

BOIKO, Viktor et al. The Optimization of the Cryptocurrency Portfolio in View of the Risks. **Journal of Management Information and Decision Sciences**, [S. l.], v. 24, p. 1-9, 2021.

ESTADÃO E-INVESTIDOR. **Investidores recorrem cada vez mais às criptomoedas, diz estudo**. Estadão, 9 ago. 2021. Disponível em: https://einvestidor.estadao.com.br/comportamento/perfil-dos-investidores-cripto-no-brasil/. Acesso em: 17 ago. 2021.

NISTRUP, Peter. **Retrieving Full Historical Data for Every Cryptocurrency on Binance & Bitmex Using the Python APIs**. Medium, 17 jul. 2019. Disponível em: https://medium.com/swlh/retrieving-full-historical-data-for-every-cryptocurrency-on-binance-bitmex-using-the-python-apis-27b47fd8137f. Acesso em: 17 ago. 2021.

FU, Tak-chung et al. Adopting genetic algorithms for technical analysis and portfolio management. **Computers & Mathematics with Applications**, [S. l.], v. 66, p. 1743-1757, 1 dez. 2013. Disponível em: https://reader.elsevier.com/reader/sd/pii/S0898122113004938. Acesso em: 17 ago. 2021.

ELTON, Edwin J et al. Modern portfolio theory, 1950 to date. **Journal of Banking & Finance**, [S. l.], v. 21, p. 1743-1759, 1 dez. 1997. Disponível em: https://www.sciencedirect.com/science/article/abs/pii/S0378426697000484. Acesso em: 17 ago. 2021.

SHARPE, William F. et al. The Sharpe Ratio. **The Journal of Portfolio Management Fall**, [S. l.], p. 49-58, 1994. Disponível em: https://jpm.pm-research.com/content/21/1/49. Acesso em: 17 ago. 2021.

DE CASTRO, Fabrício. Juro bancário médio cai em 2020, mas taxa do cartão de crédito avança**. O Estado de S.Paulo**, [S. l.], 28 jan. 2021. Disponível em: https://economia.estadao.com.br/noticias/geral,juro-bancario-medio-cai-em-2020-mas-taxa-do-cartao-de-credito-avanca,70003596966. Acesso em: 7 ago. 2021.

BOURGUIGNON, Natalia. Bitcoin em alta: entenda os riscos e como investir na moeda digital. **Rede Gazeta**, [S. l.], 10 fev. 2021. Disponível em: https://www.agazeta.com.br/es/economia/bitcoin-em-alta-entenda-os-riscos-e-como-investir-na-moeda-digital-0221. Acesso em: 7 ago. 2021.

COMO INVESTIR em Bitcoin no Brasil?: Bitcoin vai subir mais em 2021?. **Blog Binance**, 26 abr. 2021. Disponível em: https://www.binance.com/pt-BR/blog/all/como-investir-em-bitcoin-no-brasil-bitcoin-vai-subir-mais-em-2021-421499824684902017. Acesso em: 1 fev. 2022.

MALLAWAARACHCHI, Vijini. Introduction to Genetic Algorithms. **Towards Data Science**, 7 jul. 2017. Disponível em: https://towardsdatascience.com/introduction-to-genetic-algorithms-including-example-code-e396e98d8bf3. Acesso em: 1 fev. 2022.

ROULETTE WHEEL SELECTION IN PYTHON. **Roc Reguant**, 2019. Disponível em: https://rocreguant.com/roulette-wheel-selection-python/2019/. Acesso em: 1 fev. 2022.

SINGH GILL, Navdeep. **Machine learning Model Validation Testing**. XenonStack, 21 nov. 2018. Disponível em: https://www.xenonstack.com/insights/what-is-model-validation-testing. Acesso em: 3 fev. 2022.

BOUHMALA, Noureddine. A Multilevel Genetic Algorithm for the Maximum Satisfaction Problem. In: FERNANDEZ, Marco A. A. **Artificial Intelligence**: Emerging Trends and Applications. Mexico: IntechOpen, 01/11/2017. cap. 13, p. 265-274. Disponível em: https://www.intechopen.com/books/6646. Acesso em: 3 fev. 2022.