



OTIMIZAÇÃO DE PORTFÓLIO COM ATIVOS DO IBOVESPA USANDO ALGORITMOS GENÉTICOS

Daniel Mognato Vaneli
Instituto Federal do Espírito Santo
E-mail: danielmognato@hotmail.com.

Guilherme Guilhermino Neto
Instituto Federal do Espírito Santo
E-mail: guilherme.neto@ifes.edu.br.

1 INTRODUÇÃO

No contexto atual, no qual os mercados financeiros estão em constante evolução e os investidores buscam incessantemente por estratégias eficazes de alocação de ativos, a aplicação de técnicas avançadas de análise de dados torna-se uma ferramenta essencial. O mercado financeiro, como um ambiente dinâmico e altamente volátil, oferece uma vasta gama de oportunidades e desafios para aqueles que procuram otimizar suas carteiras de investimentos.

Como abordado por Assaf (2001), o mercado financeiro, como um sistema complexo, é influenciado por uma infinidade de fatores, desde indicadores macroeconômicos até eventos geopolíticos. Para lidar com essa complexidade, é crucial a utilização de ferramentas analíticas poderosas. Harry Markowitz (1952) desenvolveu a Teoria do Portfólio que tem como objetivo foi reduzir ao máximo o risco ao qual grandes retornos estão associados, isto é, maximizar o retorno e minimizar o risco da carteira analisada.

Diante disso, os algoritmos genéticos surgem como uma técnica promissora para otimizar a alocação de ativos em uma carteira de investimentos. Inspirados no processo de evolução natural, esses algoritmos se baseiam em princípios como seleção natural, recombinação genética e mutação para encontrar soluções ótimas em espaços de busca complexos, conforme caracterizado por Chen & Zhou (2021). Ao aplicar algoritmos genéticos à alocação de ativos, pode-se explorar de forma eficiente o espaço de possíveis carteiras, buscando maximizar retornos e minimizar riscos.

Além disso, este trabalho também inclui uma análise comparativa com uma estratégia de comprar o índice Ibovespa diretamente, uma abordagem mais tradicional de investimento, que envolve a simples compra e retenção de uma cesta de ativos ao longo do tempo. Essa comparação permite avaliar o desempenho relativo das estratégias propostas em termos de métricas importantes, como Índice de *Sharpe Ratio* e *Drawdown*. O índice *Sharpe Ratio* é uma medida comumente utilizada para o retorno do risco dado pelo investimento (Hodges, 1998). Com ele é possível comparar

diferentes investimentos, considerando não apenas os retornos esperados, mas também o nível de risco assumido para alcançar esses retornos.

Este trabalho se propõe a explorar e comparar os resultados da estratégia de otimização de ativos em um portfólio aplicado aos ativos do Ibovespa focando na maximização da carteira através da ponderação adequada de cada ativo. Para alcançar esse objetivo, utiliza-se uma abordagem baseada em ciência de dados, que envolve desde a aquisição e tratamento dos dados até a implementação de modelos sofisticados para tomada de decisão.

Portanto, este estudo não apenas contribui para o avanço do conhecimento na área de finanças quantitativas e ciência de dados aplicada ao mercado financeiro, mas também oferece ideias valiosas para investidores e gestores de carteiras interessados em maximizar o retorno de seus investimentos enquanto gerenciam o risco de forma eficaz.

2 METODOLOGIA

Nesta seção, será descrita a metodologia aplicada a este trabalho, desde o sorteio de ativos e aquisição dos dados até a implementação do algoritmo genético para otimização da alocação de ativos. Cada etapa é fundamental para o desenvolvimento e análise das estratégias aplicadas aos ativos do Ibovespa.

2.1 Fluxograma do Algoritmo

A Figura 1 apresenta uma simplificação da lógica por trás do algoritmo em estudo, destacando as etapas desde a seleção dos ativos até a implementação do algoritmo genético.

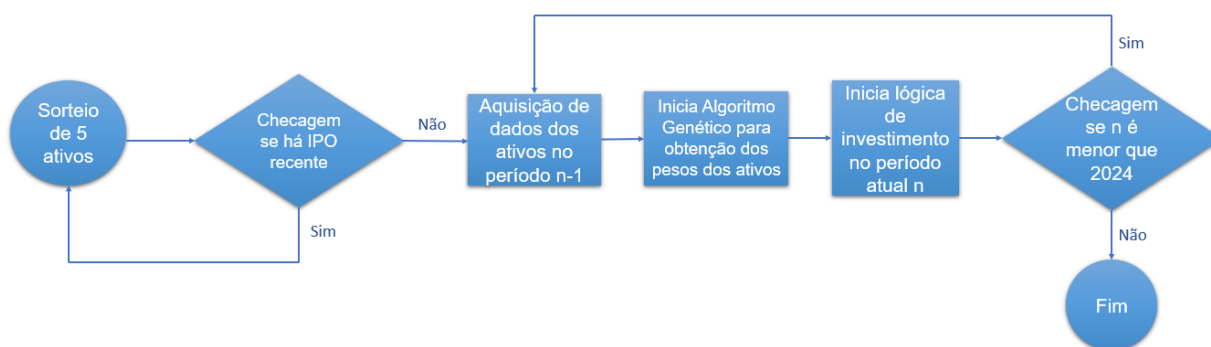


Figura 1 - Fluxograma principal do algoritmo (autor).

2.2 Sorteio de Ativos e Aquisição de Dados

A metodologia inicial se baseia no sorteio aleatório de cinco ativos dentro de uma cesta composta por ações do índice Ibovespa no ano de 2024, simulando a



criação de um portfólio aleatório. A escolha de cinco ativos visa manter a simplicidade e clareza na análise, ao mesmo tempo que demonstra os princípios de diversificação e otimização de portfólio. Este número é suficiente para distribuir o risco entre diferentes ativos, sem tornar o processo excessivamente complexo, permitindo uma avaliação detalhada de cada ativo escolhido. Após a seleção dos cinco ativos, é verificado qual deles possui o IPO (*Initial Public Offering* – Oferta Pública Inicial) mais recente para definir o ano inicial de aquisição dos dados necessários para análise. Essa verificação é essencial porque, se um dos ativos tiver IPO recente, não haverá uma base de dados suficientemente grande para o estudo.

Os dados históricos dos ativos são adquiridos através da API (*Application Programming Interface* – Interface de Programação de Aplicação) do *Yahoo Finance*, em formato de séries temporais diárias para cada ano em estudo, ou seja, para cada ano em estudo, foi extraído o preço de fechamento ajustado do dia de cada ativo durante todo ano. De todas as simulações rodadas pelo algoritmo, o período mais longo de avaliação foi de 2013 até 2024. O preço de fechamento ajustado leva em conta todos os desdobramentos e distribuições de dividendos aplicáveis, garantindo que a análise considere todos os eventos relevantes para os preços históricos.

2.3 Implementação do Algoritmo Genético

Com os dados em mãos, a próxima etapa consiste na implementação do algoritmo genético para otimização da alocação de ativos na carteira. O algoritmo é projetado para encontrar os pesos ideais para cada ativo, maximizando o retorno da carteira e minimizando o risco. A linguagem de programação escolhida foi o Python, utilizando as bibliotecas *numpy*, *pandas* e *matplotlib*.

2.3.1 Funcionamento dos Algoritmos Genéticos

Os algoritmos genéticos são uma técnica de otimização inspirada no processo de evolução natural. Eles operam em uma população de soluções candidatas, onde cada solução é representada por um conjunto de parâmetros (genes) que define uma possível alocação de ativos na carteira. O processo de otimização ocorre ao longo de várias gerações, com os indivíduos mais aptos sendo selecionados para reprodução e produção de descendentes.

O processo começa com a geração aleatória de uma população inicial de 400 soluções candidatas. Cada indivíduo representa uma alocação de ativos na carteira, com os pesos dos ativos sendo os genes do indivíduo. Para o problema em questão, um tamanho de população grande se mostrou benéfico para capturar uma variedade maior de soluções e evitar a convergência prematura para mínimos locais.



2.3.2 Avaliação e Seleção

Cada indivíduo na população é avaliado com base em uma função de aptidão, que quantifica o desempenho da alocação de ativos. Uma das métricas mais conhecidas para avaliar o desempenho de uma carteira, e que foi utilizada neste trabalho, é a razão de Sharpe. Através da razão de Sharpe, é possível formular uma carteira que maximiza a relação risco-retorno, onde o risco é mensurado pela variância dos retornos. Assim, a razão de Sharpe, definida pela Equação 1 como:

$$SR = \frac{r_m - r_{lr}}{\sigma_p} \quad (1)$$

onde r_m é o retorno médio da carteira, r_{lr} é o retorno livre de risco e, σ_p o desvio-padrão (risco) da carteira. É uma relação do retorno excedente em relação ao risco da carteira. Mostra se o retorno excedente é oriundo da estratégia da carteira ou dos riscos altos. Quanto maior o Índice Sharpe, mais atraente é o retorno ajustado ao risco. A taxa SELIC foi adotada como retorno livre de risco devido à segurança do emissor (governo), alta liquidez e proteção contra a inflação. Os dados históricos da taxa SELIC foram obtidos através da API DadosAbertosBrasil. Com isso, a carteira que maximiza a razão de Sharpe, é obtida como a solução do seguinte problema de otimização na Equação 2:

$$\sigma_p^2 = \sum_{i=1}^n \sum_{j=1}^n w_i w_j Cov(r_i, r_j) \quad (2)$$

onde n é o número de ativos, $w_i w_j$ são os pesos do i -ésimo e j -ésimo ativo na carteira e $Cov(r_i, r_j)$ é a covariância entre os ativos i e j . Dessa forma, com a mudança dos pesos $w_i w_j$ dos ativos e da covariância entre eles, busca-se minimizar a variância σ_p^2 para então maximizar a razão de *Sharpe* na Equação 1.

2.3.3 Reprodução e Mutação

Os indivíduos mais aptos são selecionados para reprodução, com probabilidade proporcional à sua aptidão. Os indivíduos selecionados são combinados para produzir descendentes por meio de operadores de cruzamento, misturando os genes dos pais para gerar novas soluções. Para este trabalho, uma taxa de cruzamento de 0,8 foi utilizada para promover a diversidade genética.

Após o cruzamento, os genes dos descendentes podem sofrer mutação com uma pequena probabilidade. Uma taxa de mutação de 0,01 foi adotada permitindo a introdução de novas informações genéticas sem mudanças drásticas nas soluções candidatas.



2.3.4 Elitismo e Convergência

Os descendentes gerados substituem parte da população anterior, garantindo uma mistura de indivíduos antigos e novos. O elitismo, com uma taxa de 0,1 foi aplicado para manter a qualidade das soluções ao longo das gerações. O processo de evolução continua por um número fixo de gerações, ao final das quais o melhor indivíduo é selecionado como a solução final.

2.4 Lógica de Investimento

Na etapa final, é realizada a lógica de investimento, onde os cinco ativos são comprados no primeiro dia útil do ano utilizando os pesos otimizados pelo algoritmo genético. A venda é realizada no último dia útil do ano. Além da razão de Sharpe, outra medida de performance utilizada é a perda máxima (*maximum drawdown*), que indica a maior perda de capital antes de uma recuperação.

2.5 Comparação com o Índice Ibovespa

Os dados históricos do índice Ibovespa foram adicionados para comparação de resultados com um benchmark de renda variável. Assim, é possível avaliar, através da razão de Sharpe e do *drawdown*, a viabilidade de comprar cinco ativos aleatórios com pesos otimizados versus simplesmente comprar o índice Ibovespa.

A metodologia descrita combina técnicas de sorteio aleatório, aquisição de dados, e algoritmos genéticos para otimização de portfólios. A implementação e análise dessas estratégias fornecem insights sobre a eficácia da otimização da alocação de ativos e sua comparação com benchmarks tradicionais.

3 RESULTADOS E DISCUSSÃO

Este capítulo expõe e discute os resultados obtidos ao longo do estudo. Foram analisadas 100 repetições do código, gerando aproximadamente 750 rodadas do algoritmo genético. Como cada carteira de 5 ativos são sempre diferentes, a quantidade de anos analisados será variável. Aqui serão apresentados alguns gráficos plotados durante a execução do algoritmo que descrevem melhor cada etapa de resultados.

Na Figura 2 é apresentado a evolução do melhor índice Sharpe por cada geração do algoritmo genético. Pode-se observar que a partir de aproximadamente 200 gerações, o índice praticamente converge para uma solução ótima. A variação após as 200 gerações não ultrapassa os 1%. Isso indica que, para aquele domínio de estudo, os parâmetros utilizados para o algoritmo genético se mostraram eficazes para a convergência do problema.

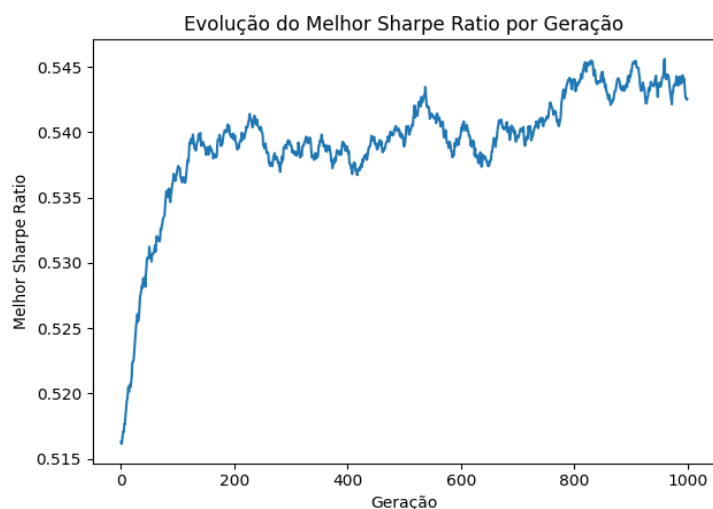


Figura 2 - Evolução da Melhor Razão Sharpe por Geração (Autor).

Dentre todos os testes rodados, a Figura 3a apresenta um exemplo de alocação ótima dentre os cinco ativos na carteira. Ainda na Figura 3b, é apresentado a matriz de covariância de Pearson mostrando a relação entre os ativos escolhidos. Nota-se uma boa relação entre os ativos CPFE3 e EQTL3, ambas do setor elétrico. Após calculado o melhor índice de Sharpe a alocação ficou definida como 12,5% de ABEV3, 21,8% de BRAP4, 31,6% de EQTL3, 16,4% de CPFE3 e 17,7% de KLB11.

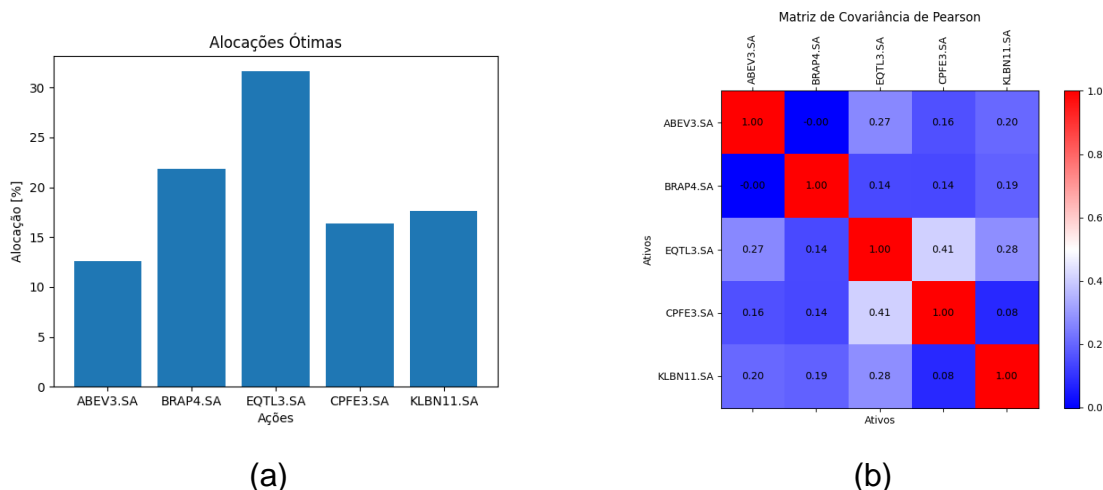


Figura 3 – (a) Alocação dos 5 Ativos na Carteira e (b) Matriz de Covariância de Pearson (Autor).

Essa combinação de pesos para os ativos advém da fronteira eficiente, conforme teoria moderna de Markowitz, que é a representação gráfica das melhores combinações possíveis de ativos que oferecem o maior retorno esperado para um

nível dado de risco (desvio-padrão) ou, de forma equivalente, o menor risco para um dado nível de retorno esperado.

A Figura 4 apresenta a fronteira eficiente para o exemplo aqui discutido, isto é, o processo do algoritmo genético aplicado à teoria de Markowitz resultou na obtenção de uma fronteira eficiente, caracterizada por carteiras que oferecem a melhor relação entre retorno esperado e risco, considerando a variabilidade e correlação dos retornos dos ativos selecionados. Essa abordagem estocástica permitiu explorar um amplo espaço de soluções possíveis, proporcionando uma otimização robusta e adaptativa da alocação de ativos.

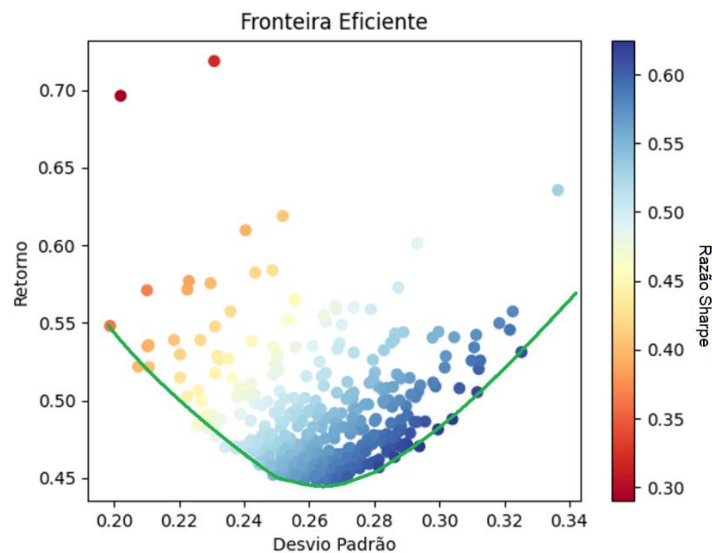


Figura 4 - Fronteira Eficiente (Autor).

Como abordado no capítulo 2, outra forma de avaliação da performance é a perda máxima ou *drawdown*. Na Figura 5 pode-se avaliar a perda máxima separada por ativo dentro da carteira otimizada. É notável a queda de aproximadamente 40% no valor dos ativos para o tempo em estudo. Isso se deu pela pandemia de COVID19 onde, em meados de março de 2020, diversos estados declararam o *lockdown*.

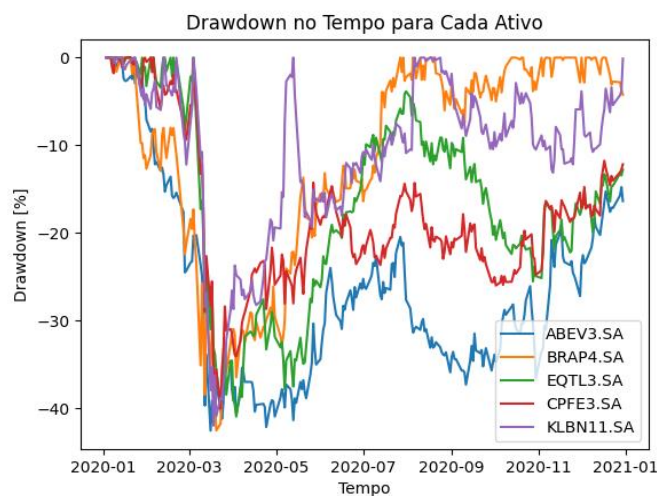


Figura 5 - Perda Máxima no Tempo para Cada Ativo da Carteira (Autor).

A Tabela 1 apresenta todos os resultados sumarizados. Os dados apresentados fornecem uma visão comparativa do desempenho de duas estratégias de investimento.

Tabela 1 - Resultados sumarizados (Autor).

	IBOVESPA	Algoritmo Genético
<i>Maior rentabilidade</i>	117%	532%
<i>Maior Prejuízo</i>	0%	-39%
<i>Operações no lucro</i>	100%	83%
<i>Operações no prejuízo</i>	0%	17%
<i>Maior Drawdown</i>	46%	70%
<i>Maior Sharpe</i>	0.73	1.49
<i>Menor Sharpe</i>	-0.65	-0.92

A estratégia do algoritmo genético demonstrou a maior rentabilidade, alcançando impressionantes 532%, enquanto a estratégia para o Ibovespa rendeu 117%. No entanto, é importante notar que a maior rentabilidade também veio com o maior prejuízo, que foi de -39% para o algoritmo genético.

Em termos de operações no lucro, o investidor que comprou Ibovespa teve um desempenho perfeito de 100% das operações no lucro. Já a estratégia do algoritmo genético, apesar de 17% das operações estarem no prejuízo, mostrou ter uma taxa relativamente boa de 83% das operações no lucro.

O índice de Sharpe, que mede o retorno ajustado ao risco, foi mais alto para o algoritmo genético, indicando um melhor desempenho ajustado ao risco em comparação com o Ibovespa. Porém, quando se é avaliado a perda máxima



(*drawdown*) nota-se que o algoritmo genético apresentou um valor bem elevado de 70% mostrando que, apesar da elevada rentabilidade, índice de Sharpe e operações no lucro, o investidor amargaria um período de queda do seu patrimônio investido na ordem de 70%.

4 CONCLUSÕES E CONSIDERAÇÕES FINAIS

Com base na análise dos dados, pode-se concluir que a estratégia do algoritmo genético proporcionou a maior rentabilidade, mas também veio com um risco significativo, como evidenciado pelo maior prejuízo e *Drawdown*. Por outro lado, a estratégia atrelada ao Ibovespa teve um desempenho consistente com 100% de lucros e nenhum prejuízo.

No entanto, ao considerar o retorno ajustado ao risco, medido pelo índice de Sharpe, o algoritmo genético superou o Ibovespa em ambas as estratégias, sugerindo que, apesar do risco adicional, o algoritmo genético pode oferecer um potencial de retorno superior quando o risco é adequadamente gerenciado.

Portanto, a escolha da estratégia de investimento e da composição do portfólio deve ser baseada não apenas na rentabilidade, mas também na tolerância ao risco do investidor. Investidores com maior tolerância ao risco podem preferir a estratégia do algoritmo genético, enquanto aqueles com menor tolerância ao risco podem achar a estratégia do Ibovespa mais adequada.

AGRADECIMENTOS

Primeiramente, meu mais sincero agradecimento ao meu orientador e professor, Guilherme. Sua orientação, expertise e apoio ao longo deste processo foram essenciais para o desenvolvimento deste trabalho.

Em segundo lugar, gostaria de expressar minha profunda gratidão à minha esposa, Thaylis. Seu amor, encorajamento e apoio emocional foram fundamentais para que eu pudesse enfrentar os desafios e superar os obstáculos ao longo desta jornada acadêmica.

A ambos, meu sincero obrigado por fazerem parte deste momento importante da minha vida e por tornarem possível a realização deste trabalho.



DECLARAÇÃO DE USO DE TECNOLOGIAS AUXILIADAS POR INTELIGÊNCIA ARTIFICIAL

O ChatGPT foi a única tecnologia de inteligência artificial utilizada neste trabalho. No contexto deste estudo, o ChatGPT foi utilizado para revisar e aprimorar a redação do texto, fornecendo sugestões para melhorar a coesão, clareza e fluidez do conteúdo. Além disso, o ChatGPT também desempenhou um papel importante no auxílio ao tratamento de erros encontrados no código Python desenvolvido para a pesquisa, fornecendo insights e sugestões para correções.

REFERÊNCIAS

Assaf Neto, Alexandre. **Mercado Financeiro**. 4. ed. São Paulo: Atlas, 2001.

Chen, S. & Zhou, C (2021). “**Stock Prediction Based on Genetic Algorithm Feature Selection and Long Short-Term Memory Neural Network**,” IEEE Access, 9, 9066-9072, 10.1109/ACCESS.2020.3047109.

DadosAbertosBrasil. Documentação da biblioteca. Disponível em: <
<https://www.gustavofurtado.com/dab.html>> Acesso: 13 mai. 2024.

Hodges, S. D. **A generalization of the Sharpe Ratio and its applications to valuation bounds and risk measures**. 1998.

Markowitz, H. **Portfolio Selection**. Journal of Finance. 7^a ed. p.77-91. 1952.

Daniel Mognato Vaneli

OTIMIZAÇÃO DE PORTIFÓLIO COM ATIVOS DO IBOVESPA USANDO ALGORITMOS GENÉTICOS

Trabalho Final de Curso apresentado à
Coordenadoria do Curso de Engenharia de
Produção do Instituto Federal do Espírito Santo
– *campus* Cariacica como requisito parcial para
obtenção do título de Especialista em
Engenharia de Produção com Ênfase em
Ciência de Dados

Aprovado em 24 de junho de 2024

COMISSÃO EXAMINADORA

Prof. Guilherme Guilhermino Neto, D.Sc.

Ifes – Instituto Federal do Espírito Santo

Orientador

Prof. Tiago José Menezes Gonçalves, D.Sc.

Ifes – Instituto Federal do Espírito Santo

Membro da banca avaliadora

Prof. Erivelto Fioresi de Sousa, D.Sc.

Ifes – Instituto Federal do Espírito Santo

Membro da banca avaliadora



FOLHA DE APROVAÇÃO-TCC Nº 16/2024 - CAR-CCEP (11.02.19.01.08.03.10)

(Nº do Protocolo: NÃO PROTOCOLADO)

(Assinado digitalmente em 30/07/2024 17:01)

ERIVELTO FIORESI DE SOUSA

PROFESSOR DO ENSINO BASICO TECNICO E TECNOLOGICO

CAR-CCE (11.02.19.01.08.03.11)

Matrícula: 1579284

(Assinado digitalmente em 30/07/2024 16:55)

GUILHERME GUILHERMINO NETO

PROFESSOR DO ENSINO BASICO TECNICO E TECNOLOGICO

CAR-CCEP (11.02.19.01.08.03.10)

Matrícula: 2151589

(Assinado digitalmente em 31/07/2024 09:43)

TIAGO JOSE MENEZES GONCALVES

PROFESSOR DO ENSINO BASICO TECNICO E TECNOLOGICO

CAR-CCEP (11.02.19.01.08.03.10)

Matrícula: 2073974

Visualize o documento original em <https://sipac.ifes.edu.br/documentos/> informando seu número: **16**, ano: **2024**,
tipo: **FOLHA DE APROVAÇÃO-TCC**, data de emissão: **30/07/2024** e o código de verificação: **d7c0223f2c**