

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PARANÁ
ESCOLA POLITÉCNICA
PROGRAMA DE PÓS GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO E SISTEMAS

AMANDA TROJAN FENERICH

MODELO BASEADO EM TÉCNICAS DE APRENDIZAGEM DE MÁQUINA E
ALGORITMO GENÉTICO PARA ANÁLISE DE DECISÃO MULTICRITÉRIO EM
PROBLEMAS DE SELEÇÃO DE PORTFÓLIO

CURITIBA - PR

2018

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PARANÁ
ESCOLA POLITÉCNICA
PROGRAMA DE PÓS GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO E SISTEMAS

AMANDA TROJAN FENERICH

**MODELO BASEADO EM TÉCNICAS DE APRENDIZAGEM DE MÁQUINA E
ALGORITMO GENÉTICO PARA ANÁLISE DE DECISÃO MULTICRITÉRIO EM
PROBLEMAS DE SELEÇÃO DE PORTFÓLIO**

Tese apresentada ao Programa de Pós-Graduação em Engenharia de Produção e Sistemas, da Pontifícia Universidade Católica do Paraná, como requisito para obtenção do título de Doutor (a) em Engenharia de Produção e Sistemas.

Área de Concentração: Gerência de Produção e Logística

Orientadora: Profa. PhD.Eng. Maria Teresinha Arns Steiner

CURITIBA - PR

2018

Dados da Catalogação na Publicação
Pontifícia Universidade Católica do Paraná
Sistema Integrado de Bibliotecas – SIBI/PUCPR
Biblioteca Central

F332m
2018 Fenerich, Amanda Trojan
Modelo baseado em técnicas de aprendizagem de máquina e algoritmo genético para análise de decisão multicritério em problemas de seleção de portfólio / Amanda Trojan Fenerich ; orientadora: Maria Teresinha Arns Steiner. – 2018.
145 f. : il. ; 30 cm

Tese (doutorado) – Pontifícia Universidade Católica do Paraná, Curitiba, 2018
Bibliografia: f. 112-126

1. Engenharia de produção. 2. Processo decisório. 3. Algoritmos genéticos. 4. Otimização combinatória. I. Steiner, Maria Teresinha Arns. II. Pontifícia Universidade Católica do Paraná. Programa de Pós-Graduação em Engenharia de Produção e Sistemas. III. Título.

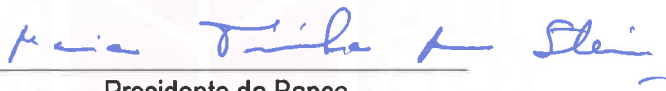
CDD 20. ed. – 670

TERMO DE APROVAÇÃO

Amanda Trojan Fenerich

“MODELO BASEADO EM TÉCNICAS DE APRENDIZAGEM DE MÁQUINA E ALGORITMO GENÉTICO PARA ANÁLISE DE DECISÃO MULTICRITÉRIO EM PROBLEMAS DE SELEÇÃO DE PORTFÓLIO”.

Tese aprovada como requisito parcial para obtenção do grau de Doutor no Curso de Doutorado em Engenharia de Produção e Sistemas, Programa de Pós-Graduação em Engenharia de Produção e Sistemas, da Escola Politécnica da Pontifícia Universidade Católica do Paraná, pela seguinte banca examinadora:



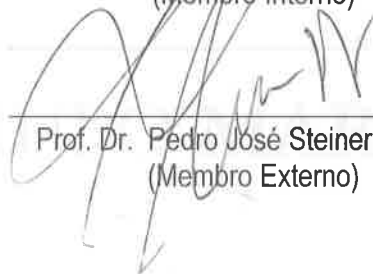
Presidente da Banca
Prof.ª Dra. Maria Teresinha Arns Steiner
(Orientadora)



Prof. Dr. Eduardo de Freitas Rocha Loures
(Membro Interno)



Prof. Dr. Osiris Canciglieri Júnior
(Membro Interno)



Prof. Dr. Pedro José Steiner Neto
(Membro Externo)



Prof.ª Dra. Deise Maria Bertholdi da Costa
(Membro Externo)

Curitiba, 29 de novembro de 2018.

AGRADECIMENTOS

Em primeiro lugar, agradeço a Deus pelo maior presente que eu poderia receber: a minha Família. Agradeço minhas irmãs, Aline Trojan Fenerich e Adriana Trojan Fenerich, pelo amor e carinho incondicional, e agradeço aos meus pais, Valdemar Fenerich e Marly Trojan Fenerich, pelo apoio e oportunidade para realizar meu sonho, por acreditarem no meu potencial e por demonstrarem sempre orgulho das minhas ações e decisões.

Agradeço minha Orientadora Profa. Maria Teresinha Arns Steiner, primeiramente por aceitar a minha orientação, e também por auxiliar na elaboração desta tese. E a agradeço principalmente pela instrução, paciência e compreensão em todas as situações complicadas no decorrer desta etapa tão importante de minha vida.

A todos os Professores do Programa de Pós-Graduação em Engenharia de Produção e Sistemas da PUCPR, e demais professores de programas vinculados, que ao longo destes quatro anos de doutorado, me ensinaram lições valiosas. Também agradeço ao pessoal da secretaria e toda a equipe do PPGEPS/PUCPR, que sempre estiveram à disposição no que foi preciso para me auxiliar.

Aos meus amigos que conheci no passado e durante o curso, que conviveram comigo durante dias e noites duras de estudos, nem sempre agradáveis, mas com certeza todo esse tempo de convívio proporcionou momentos de alegrias que ficarão para sempre na memória e no coração. E, principalmente, agradeço meu amigo e companheiro, Diego Paolo Tsutsumi, que me acompanhou, apoiou e me deu forças para continuar e finalizar esta etapa.

E por último, mas não menos importante, agradeço a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo apoio financeiro concedido durante todo o período do curso, para que pudesse desenvolver as atividades de pesquisa.

A todos, o meu muito obrigado!!!

RESUMO

A pesquisa trata sobre o problema de seleção de portfólio de abordagem multicritério, com a integração dos principais algoritmos classificadores do campo de aprendizagem de máquina, juntamente com os algoritmos genéticos. Alguns métodos de decisão multicritério, incluindo os métodos para a problemática de portfólio, exigem a definição de parâmetros, muitas vezes obtidos pela elicitación junto à decisores. No caso do modelo aditivo para seleção de portfólio abordado nesta pesquisa, os parâmetros necessários consistem nas constantes de escala (pesos) de cada um dos critérios/atributos considerados na análise. Porém, a definição destes parâmetros é uma tarefa difícil de ser interpretada e executada, mesmo até por especialistas da área. Também, em situações recorrentes de tomada de decisão, formam-se um montante de informações, que podem ser utilizadas para auxiliar e melhorar o processo decisório. Sendo assim, o objetivo desta pesquisa é propor um modelo de análise para problemas de seleção de portfólio no contexto de decisão multicritério, incorporando técnicas de aprendizagem de máquina e de algoritmos genéticos, para determinar e otimizar as constantes de escala (pesos) dos critérios. O modelo proposto foi incorporado em *framework*, constituído de quatro fases, cada qual com um objetivo principal: a) Fase Preliminar, para caracterização geral do problema; b) Fase de predição, para o reconhecimento de padrões de decisões históricas e determinação da ordem dos critérios; c) Fase de Elicitación com o decisor: para estabelecer a ordem dos critérios, com base nas preferências do decisor; e d) Fase de Seleção e Otimização: para selecionar os projetos que irão compor o portfólio, e otimização dos valores das constantes de escala (pesos). Esta última fase foi implementada em um *software* de apoio a decisão, e o modelo foi testado com a aplicação de um exemplo numérico, e também com dois estudos de caso: seleção de projetos em empresa de energia elétrica; e seleção de clientes para concessão de crédito bancário. Em ambos os estudos de casos, a otimização das constantes de escala (pesos), obtida a partir dos algoritmos genéticos, resultou em melhores valores de portfólio ótimo, sugerindo assim que o processo de elicitación com o decisor pode ser facilitado ao utilizar o modelo proposto. Ou seja, mesmo quando o decisor não for capaz de informar suas preferências para definir as constantes de escala, o modelo pode ser empregado para definição de um resultado satisfatório para o problema.

Palavras-chave: Sistema de Apoio à Decisão; Seleção de Portfólio; Algoritmos Classificadores; Seleção de Atributos; Otimização Combinatória.

ABSTRACT

The research deals with the problem of portfolio selection in multicriteria approach, with the integration of the main classifier algorithms of the machine learning field, along with the genetic algorithms. Some multicriteria decision methods, including the methods for the portfolio problem, require the definition of parameters, often obtained by elicitation with decision makers. In the case of the additive model for portfolio selection addressed in this research, the necessary parameters consist of the scale constants (weights) of each of the criteria/attributes considered in the analysis. However, the definition of these parameters is a difficult task to be interpreted and performed, even by specialists in the field. Also, in recurrent decision-making situations, there is an amount of information that can be used to aid and improve the decision-making process. Thus, the objective of this research is to propose an analysis model for portfolio selection problems in the multicriteria decision context, incorporating machine learning techniques and genetic algorithms, to determine and optimize the scale constants (weights) of the criteria. The proposed model was incorporated in a framework, consisting of four phases, each with a main objective: a) Preliminary Phase, for general characterization of the problem; b) Prediction phase, for the recognition of patterns of historical decisions and determination of the order of the criteria; c) Elicitation phase with the decision maker: to establish the order of the criteria, based on the preferences of the decision maker; and d) Selection and Optimization Phase: to select the projects that will compose the portfolio, and optimization of the values of the scale constants (weights). This last phase was implemented in a decision support software, and the model was tested with the application of a numerical example, and also with two case studies: project selection in an electric power company; and selection of clients for granting bank credit. In both case studies, the optimization of the scale constants (weights), obtained from the genetic algorithms, resulted in better optimal portfolio values, thus suggesting that the elicitation process with the decision maker can be facilitated by using the proposed model. That is, even when the decision maker is not able to inform his preferences to define the scale constants, the model can be used to define a satisfactory result for the problem.

Keywords: Decision Support System; Portfolio Selection; Classification Algorithms; Feature Selection; Combinatorics Optimization.

LISTA DE FIGURAS

Figura 1.1 - Classificação da Pesquisa	20
Figura 2.1 - Exemplo de uma rede bayesiana aplicada ao problema de compra de roupa.....	25
Figura 2.2 - Rede Neural Artificial, <i>Multilayer Perceptron</i>	26
Figura 2.3 - Árvore de Decisão, exemplo para compra de carro.....	27
Figura 2.4 - Esquema da Máquina de Vetores de Suporte (<i>Support Vector Machine, SVM</i>).29	
Figura 2.5 - Curva ROC (<i>Receiver Operating Characteristic</i>)	31
Figura 2.6 - Operações de Cruzamento e Mutação	34
Figura 3.1 - <i>Framework</i> do processo de revisão sistemática, análise bibliométrica e análise de conteúdo	35
Figura 3.2 - Quantidade de documentos encontrados nas bases de dados com diferentes algoritmos de busca	38
Figura 3.3 - Quantidade de documentos encontrados por ano nas bases de dados	40
Figura 3.4 - Número de citações por ano dos documentos nas bases de dados	40
Figura 3.5 - Evolução dos sete documentos mais citados nos últimos 10 anos, nas bases de dados Scopus	42
Figura 3.6 - Autores com maior número de documentos, nas bases de dados	42
Figura 3.7 - Instituições com maior número de documentos, nas bases de dados	43
Figura 3.8 - Países com maior número de documentos, nas bases de dados.....	44
Figura 3.9 - Periódicos com maior número de documentos, nas bases de dados.....	44
Figura 3.10 - Áreas do conhecimento das bases mais significativas.....	45
Figura 3.11 - Nuvens de palavras com títulos e palavras-chave dos documentos encontrados	45
Figura 3.12 - Conceitos e abordagens elencados nos documentos encontrados	46
Figura 3.13 - Principais áreas dos métodos utilizados nos documentos.....	47
Figura 3.14 - Número de documentos que empregaram os métodos relacionados a Metaheurísticas.....	48
Figura 3.15 - Número de documentos que empregaram os métodos relacionados a Decisão Multicritério.....	50
Figura 3.16 - Número de documentos que empregaram os métodos relacionados a Aprendizagem de Máquina.....	51
Figura 3.17 - Número de documentos que empregaram outros métodos.....	52

Figura 3.18 - Campos de aplicações práticas nos quais foram realizados estudos.....	53
Figura 3.19 - Proporção dos tipos de pesquisa nos quais se enquadram os trabalhos correlatos	56
Figura 4.1 - <i>Framework</i> geral do modelo proposto.....	63
Figura 4.2 - Sequência do fluxo de informações para o caso de existir registros históricos e o decisor estiver disponível	66
Figura 4.3 - Sequência do fluxo no processo decisório para o caso em que não tem registros históricos nem a disponibilidade do decisor.....	68
Figura 4.4 - Detalhes dos Procedimentos da Fase de Predição	71
Figura 4.5 - Detalhes dos procedimentos da Fase de Elicitação com Decisor	73
Figura 4.6 - Detalhes dos procedimentos da Fase de Seleção e Otimização.....	75
Figura 5.1 - Interface gráfica do procedimento de Otimização Combinatória.....	80
Figura 5.2 - Interface gráfica do procedimento de Computação Evolucionária.....	81
Figura 5.3 - Interface gráfica das informações mostradas com os botões “ <i>Show more information</i> ”	82
Figura 5.4 - Interface com o resultado obtido para o vetor <i>KED</i> do problema hipotético.....	83
Figura 5.5 - Interface com o resultado da otimização pelo algoritmo genético	84
Figura 5.6 - Evolução de cada iteração do procedimento de otimização pelo algoritmo genético.....	85
Figura 5.7 - Interface parcial com o resultado das soluções encontradas para o problema hipotético	86
Figura 6.1 - Fluxo do modelo proposto para o Caso I: Seleção de projetos para empresa do setor elétrico	89
Figura 6.2 - Fluxo do modelo proposto para o Caso II: Avaliação de risco de crédito para pessoa jurídica	96

LISTA DE ABREVIATURAS E SIGLAS

ABEPRO	Associação Brasileira de Engenharia de Produção
ACO	<i>Ant Colony Optimization</i>
AD	Árvore de Decisão
AHP	<i>Analytic Hierarchy Process</i>
ANP	<i>Analytic Network Process</i>
BCA	<i>Bee Colony Algorithm</i>
BNDES	Banco Nacional de Desenvolvimento
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CPD	<i>Conditional Probability Distribution</i>
DAG	<i>Directed Acyclic Graph</i>
DEA	<i>Data Envelopment Analysis</i> ,
DEMATEL	<i>Decision Making Trial and Evaluation Laboratory</i>
DSS	Decision Support System
ELECTRE	<i>Elimination Et Choix Traduisant la Réalité</i>
GA	<i>Genetic Algorithm</i>
IA	Inteligência Artificial
INFORMS	<i>Institute for Operations Research and the Management Sciences</i>
MAUT	<i>Multi-attribute Utility Theory</i>
MCDA	<i>Multicriteria Decision Aid/Analysis</i>
MCDM	<i>Multicriteria Decision Making/Methods</i>
ML	<i>Machine Learning</i>
MVS	Máquina de Vetores de Suporte
NSGA II	<i>Non-dominated Sorting Genetic Algorithm</i>
OR	Operation Research
P&D	Pesquisa e Desenvolvimento
PO	Pesquisa Operacional
PPGEPS	Programa de Pós-Graduação em Engenharia de Produção e Sistemas
PROMETHEE	<i>Preference Ranking Organization Method for Enrichment of Evaluation</i>
PSO	<i>Particle Swarm Optimization</i>
PUCPR	Pontifícia Universidade Católica do Paraná
RB	Redes Bayesianas
RELC	<i>Risk Endurable Level of Company</i>
RL	Regressão Logística
RNA	Redes Neurais Artificiais
ROC	<i>Receiver Operating Characteristic</i>
RVNS	<i>Reduced Variable Neighborhood Search</i>
SI	Sistema de Informação
SMARTS	<i>Simple Multi-Attribute Rating Technique</i>
SVM	<i>Support Vector Machines</i>
TI	Tecnologia de Infomração
TOPSIS	<i>Technique for Order of Preference by Similarity to Ideal Solution</i>
UCI	<i>University of California Irvine (Machine Learning Repository)</i>
UTA/UTADIS	<i>Utilités Additives/Utilites Additives Discriminantes</i>
WEKA	<i>Waikato Environment for Knowledge Analysis</i>

LISTA DE QUADROS

Quadro 2.1 - Matriz de Confusão (2x2)	30
Quadro 4.1 - Esquema da Matriz de Avaliação (MA).	61
Quadro 4.2 - Esquema da Matriz de Avaliação de Dados Históricos (<i>MAh</i>)	61
Quadro 4.3 - Esquema da Matriz de Avaliação modificada (+ <i>cm</i> + 1)	64
Quadro 4.4 - Esquema da Matriz de Avaliação modificada ($- a_{ii} = 0$)	64
Quadro 4.5 - Pseudocódigo do Algoritmo de Aproximação <i>Greedy</i> para seleção do portfólio	76
Quadro 4.6 - Pseudocódigo do Algoritmo Genético <i>Steady-State</i> para otimização dos vetores de constante de escala (pesos)	77
Quadro 5.1 - Matriz de Avaliação inicial (MA) para o problema hipotético.....	83

LISTA DE TABELAS

Tabela 3.1 - Ranking dos sete documentos mais citados nas bases de dados Scopus e Web of Science.....	41
Tabela 3.2 - Autores que utilizaram técnicas/métodos de aprendizagem de máquina nos estudos realizados	51
Tabela 3.3 - Autores que realizaram estudos aplicados em diferentes empresas ou ramo de atividades.....	54
Tabela 4.1 - Resumo do objetivo de cada procedimento em cada fase do modelo proposto ..	62
Tabela 4.2 - Possíveis combinações para obtenção do resultado	67
Tabela 6.1 - Critérios e seus respectivos níveis e função de valor para o Caso I.....	91
Tabela 6.2 - Melhor vetor <i>KAG – ED</i> encontrado, e estatísticas básicas das 10 repetições feitas do AG.....	93
Tabela 6.3 - Melhor vetor <i>KAG</i> encontrado, e estatísticas básicas das 10 repetições feitas do AG	94
Tabela 6.4 - Soluções encontradas para o Caso I	95
Tabela 6.5 - Critérios e seus respectivos níveis e função de valor para o Caso II.....	99
Tabela 6.6 - Resultado das métricas de desempenho dos cinco algoritmos classificadores..	100
Tabela 6.7 - Resultado do <i>Feature Selection</i> para composição do vetor de constante de escala, <i>KFS</i>	102
Tabela 6.8 - Resultado da elicitação com o Decisor para composição do vetor de constante de escala, <i>KED</i>	103
Tabela 6.9 - Resultado da elicitação com o Decisor para composição do vetor de constante de escala, <i>KED</i> , com a adição de mais um critério (<i>c8</i> : Predição)	104
Tabela 6.10 - Melhores valores encontrados e estatísticas básicas, em 10 repetições feitas do AG com os vetores de constante de escala combinados com as matrizes de avaliação	106
Tabela 6.11 - Soluções encontradas para o Caso II.....	107

SUMÁRIO

1. INTRODUÇÃO	14
1.1. CONTEXTUALIZAÇÃO DA PESQUISA	14
1.2. DELIMITAÇÃO DO TEMA E QUESTÕES DE PESQUISA.....	15
1.3. OBJETIVOS	17
1.3.1. Objetivo Geral	17
1.3.2. Objetivos Específicos	17
1.4. RELEVÂNCIA E ORIGINALIDADE	17
1.5. COMPLEXIDADE E LIMITAÇÕES DA PESQUISA.....	19
1.6. CLASSIFICAÇÃO DA PESQUISA	19
1.7. ESTRUTURA DO DOCUMENTO	20
2. REFERENCIAL TEÓRICO	21
2.1. ANÁLISE DE DECISÃO MULTICRITÉRIO PARA SELEÇÃO DE PORTFÓLIO ...	21
2.2. APRENDIZAGEM DE MÁQUINA SUPERVISIONADA	24
2.2.1 Algoritmos classificadores: Redes Bayesianas	24
2.2.2 Algoritmos classificadores: Redes Neurais Artificiais	26
2.2.3 Algoritmos classificadores: Árvores de Decisão	27
2.2.4 Algoritmos classificadores: Regressão Logística	28
2.2.5 Algoritmos classificadores: Máquina de Vetores de Suporte	29
2.2.6 Métricas de avaliação dos algoritmos classificadores	30
2.2.7 Seleção de Atributos: ganho de informação	32
2.3. COMPUTAÇÃO EVOLUCIONÁRIA.....	32
2.3.1 Algoritmo Genético	33
3. REVISÃO DE LITERATURA E TRABALHOS CORRELATOS	35
3.1. ANÁLISE BIBLIOMÉTRICA	39
3.2. IDENTIFICAÇÃO DE CONCEITOS, MÉTODOS E APLICAÇÕES.....	46
3.3. DESCRIÇÃO DOS TRABALHOS CORRELATOS	55

4. PROPOSTA DE UM MODELO PARA ANÁLISE DE DECISÃO MULTICRITÉRIO EM PROBLEMAS DE SELEÇÃO DE PORTFÓLIO	59
4.1. DEFINIÇÃO DOS CONJUNTOS DE DADOS, VETORES E MATRIZES	59
4.2. <i>FRAMEWORK</i> DO MODELO PROPOSTO	62
4.3. DESCRIÇÃO DA FASE PRELIMINAR	69
4.4. DESCRIÇÃO DA FASE DE PREDIÇÃO	70
4.5. DESCRIÇÃO DA FASE DE ELICITAÇÃO COM DECISOR	72
4.6. DESCRIÇÃO DA FASE DE SELEÇÃO E OTIMIZAÇÃO	74
5. IMPLEMENTAÇÃO DA FASE DE SELEÇÃO E OTIMIZAÇÃO	79
5.1. DEMONSTRAÇÃO DE UM PROBLEMA HIPOTÉTICO	82
6. APLICAÇÕES DO MODELO: ESTUDOS DE CASOS	87
6.1. SELEÇÃO DE PROJETOS PARA EMPRESA DO SETOR ELÉTRICO	89
6.1.1. Caso I: Fase Preliminar	89
6.1.2. Caso I: Matriz de Avaliação Inicial.....	92
6.1.3. Caso I: Fase de Elicitação com o Decisor.....	92
6.1.4. Caso I: Fase de Seleção e Otimização.....	92
6.1.5. Caso I: Resultados.....	94
6.2. AVALIAÇÃO DE RISCO DE CRÉDITO PARA PESSOA JURÍDICA	95
6.2.1. Caso II: Fase Preliminar	96
6.2.2. Caso II: Matriz de Avaliação Inicial e de registros históricos	99
6.2.3. Caso II: Fase de Predição.....	100
6.2.4. Caso II: Fase de Elicitação com o decisor.....	102
6.2.5. Caso II: Fase de Seleção e Otimização	104
6.2.6. Caso II: Resultados	107
7. CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS	110
REFERÊNCIAS	112
ANEXO A – Exemplo do arquivo de importação dos dados.....	127

ANEXO B – Exemplo do arquivo de saída do algoritmo genético.....	128
ANEXO C – Dados do problema do estudo de Caso I.....	129
ANEXO D – Dados parciais do problema do estudo de Caso II.....	130
ANEXO E – Matriz de avaliação inicial para o Caso I.....	131
ANEXO F – Resultados das 10 repetições da aplicação do algoritmo genético para o KAG-ED - Caso I.....	132
ANEXO G – Resultados das 10 repetições da aplicação do algoritmo genético para o KAG - Caso I.....	133
ANEXO H – Matriz de avaliação inicial para o Caso II.....	134
ANEXO I – Matriz de registros históricos (parcial) do Caso II.....	135
ANEXO J – Matriz de avaliação modificada ($- a_{i y_i} = 0$) para o Caso II.....	136
ANEXO K – Matriz de avaliação modificada ($+ c_m + 1$) para o Caso II.....	137
ANEXO L – Resultados das 10 repetições da aplicação do algoritmo genético (MA e KAG-FS - Caso II).....	138
ANEXO M – Resultados das 10 repetições da aplicação do algoritmo genético (MA e KAG-ED - Caso II).....	139
ANEXO N – Resultados das 10 repetições da aplicação do algoritmo genético (MA e KAG - Caso II).....	140
ANEXO O – Resultados das 10 repetições da aplicação do algoritmo genético (MA($- a_{i y_i} = 0$) e KAG-FS - Caso II).....	141
ANEXO P – Resultados das 10 repetições da aplicação do algoritmo genético (MA($- a_{i y_i} = 0$) e KAG-ED - Caso II).....	142
ANEXO Q – Resultados das 10 repetições da aplicação do algoritmo genético (MA($- a_{i y_i} = 0$) e KAG- Caso II).....	143
ANEXO R – Resultados das 10 repetições da aplicação do algoritmo genético (MA($+ c_m + 1$) e KAG-ED - Caso II).....	144
ANEXO S – Resultados das 10 repetições da aplicação do algoritmo genético (MA($+ c_m + 1$) e KAG - Caso II).....	145

1. INTRODUÇÃO

Neste capítulo são tratadas os temas da pesquisa, tais como contextualização, problematização, objetivos e justificativas.

1.1. CONTEXTUALIZAÇÃO DA PESQUISA

A tomada de decisão é uma atividade diária em qualquer organização, fundamental para o planejamento estratégico e tático, cujo resultado impacta diretamente em seu desempenho. O processo decisório envolve, segundo Osei-Bryson e Barclay (2015), “a definição do problema, a identificação dos critérios, a ponderação dos critérios, a geração de alternativas, a classificação de cada alternativa e a computação da decisão ideal”.

Alguns fatores afetam diretamente o processo decisório, tais como o contexto em que uma decisão é tomada; a percepção, julgamento e entendimento do decisor; as consequências da decisão; a disponibilidade de recursos e ferramentas para tomada da decisão; entre outros. Ou seja, quanto mais complexa for a decisão, ou quanto mais impactantes forem as suas consequências, maiores serão os esforços necessários para obter uma boa solução.

Diante disso, a Pesquisa Operacional (PO) é área que lida com a aplicação de métodos analíticos avançados para ajudar a tomar as melhores decisões. Possui caráter multidisciplinar para o desenvolvimento de técnicas, e o emprego destas em problemas complexos proporciona soluções ótimas ou quase ótimas. Devido à natureza computacional e estatística da maioria das técnicas, a PO também possui fortes laços com a ciência da computação e ciência de dados (*Institute for Operations Research and the Management Sciences; INFORMS*).

Devido a sua importância no desenvolvimento de técnicas de apoio à decisão, a PO foi introduzida como uma das subáreas da Engenharia de Produção, definidas pela ABEPRO (Associação Brasileira de Engenharia de Produção, 2008), sendo responsável pela “resolução de problemas reais envolvendo situações de tomada de decisão, através de modelos matemáticos habitualmente processados computacionalmente”.

Alguns dos principais campos teóricos nos quais as técnicas da PO se apoiam são: mineração de dados; aprendizagem de máquina; análise de decisão; modelagem matemática; engenharia; teoria dos jogos; otimização; probabilidade e estatística; logística; redes sociais; gestão de projetos; entre outras. Neste trabalho, especial atenção será dada ao campo referente

à análise de decisão multicritério (*Multicriteria Decision Making*; MCDM); aprendizagem de máquina (*Machine Learning*; ML) e computação evolucionária, mais especificamente os algoritmos genéticos (*Genetic Algorithm*, GA).

Além disso, considerando o contexto de análise de decisão multicritério, neste trabalho será abordado exclusivamente o problema de seleção de portfólio. O problema de seleção de portfólio consiste basicamente na escolha de um subconjunto de projetos, dentre as várias opções de projetos disponíveis a serem realizados, de forma a maximizar o valor do portfólio e considerando que nem todos os projetos poderão ser contemplados.

Por exemplo, pode ser considerado um problema de seleção de portfólio a escolha de alunos a serem contemplados com uma bolsa de estudos em um curso de pós-graduação de uma instituição de ensino superior. Como o orçamento disponível do programa pode não ser suficiente para atender a todos os alunos, deverão ser selecionados apenas alguns dos alunos, sendo que essa escolha pode ser conduzida, por exemplo, por meio do modelo de programação linear inteira e binária (seleção de portfólio).

Outro exemplo similar de um problema de seleção de portfólio é o chamado problema da mochila, bastante conhecido dentro da área de PO. No problema da mochila, vários itens podem ser levados em uma viagem, porém há uma limitação de espaço da mochila. Então, serão selecionados os itens a serem levados dentro da capacidade da mochila, de tal forma a minimizar o peso da mochila, ou então maximizar o valor dos itens inseridos.

Os problemas reais mais comumente conhecidos de seleção de portfólio são aqueles relacionados à escolha de ações da bolsa de valores nas quais deverá ser investida uma determinada verba disponível de tal forma que o retorno sobre o investimento seja maximizado. Na seção 2.1 deste trabalho será apresentado o modelo formal para o problema de seleção de portfólio no contexto de decisão multicritério.

1.2. DELIMITAÇÃO DO TEMA E QUESTÕES DE PESQUISA

São inúmeras as aplicações das técnicas de PO e, especificamente na área de aprendizagem de máquina, Witten e Frank (2005) citam relevantes contribuições para os campos de aplicações em decisões de julgamento; diagnóstico médico; previsões de vendas e mercado, entre outras. Os autores mencionam que os procedimentos que utilizam ML são atraentes, pois melhoram a taxa de sucesso em decisões diversas, além de serem usados para explicar aos interessados as razões subjacentes à decisão. Similarmente, na área de análise de

decisão multicritério e no campo da computação evolucionária, os métodos analíticos desenvolvidos, e que são aplicados em diversas situações, também proporcionam melhoras nas taxas de sucesso das decisões, além de servirem como um alicerce para os decisores em problemas complexos.

Mas, para obter os benefícios descritos acima, é necessária a correta aplicação dos métodos, com a definição de todos os parâmetros e dados necessários. Por exemplo, em vários dos métodos multicritérios é necessário, além de um conjunto completo de informações de desempenho de alternativas, que o decisor informe suas preferências de maneira coerente, garantindo assim que os parâmetros, tais como limiares e pesos de critérios, reflitam sua estrutura de preferência para obter um resultado satisfatório.

Porém, em alguns casos, não é possível obter todas as informações e parâmetros necessários para executar o método, por motivos diversos: o decisor pode não estar disponível para participar do processo decisório; o decisor pode não estar totalmente confortável em definir os parâmetros, ou inseguro e indeciso para defini-los; as informações podem não estar disponíveis, ou então disponíveis apenas parcialmente. Enfim, independente do motivo, o analista precisa, de algum modo, definir estes parâmetros.

Sendo assim, surge a primeira questão de pesquisa: *Como definir as constantes de escala (pesos) de cada critério quando o decisor não estiver disponível, ou quando deseja fornecer somente a ordem de preferência dos critérios, sem realizar a elicitação de magnitude dos pesos?*

Nas situações em que a tomada de decisão é feita de forma recorrente, e com um contexto similar, não é necessária a presença constante do decisor para avaliar cada decisão, podendo este estipular padrões de preferências, e a decisão ser tomada com base nestes padrões. Este é o caso, por exemplo, das decisões para concessão de crédito bancário, em que a presença do gerente não é necessariamente obrigatória para a decisão de aprovar ou não o crédito para cada cliente que o solicita, uma vez que existem padrões de referência que estipulam o limite a ser aprovado, dependendo das características do cliente. E é nestes tipos de decisões que surge a segunda questão de pesquisa: *Como as informações de decisões passadas (registros históricos de decisões recorrentes) podem ser utilizadas para auxiliar na tomada das novas decisões?*

Uma terceira questão de pesquisa surge quando considera-se resolver as duas questões anteriores, ou seja: *Como trabalhar com estas duas questões (definir os pesos e dados históricos) de uma forma integrada?*

Para tentar solucionar esta questão, a pesquisa apresenta um modelo híbrido, baseando-se em métodos de MCDM, ML e AG, mais especificamente no método multicritério para problema de seleção de portfólio, nos algoritmos classificadores de aprendizagem de máquina supervisionada, e no algoritmo genético do tipo *steady-state*.

1.3. OBJETIVOS

1.3.1. Objetivo Geral

Propor um modelo de análise (ou sistema de apoio à decisão, *Decision Support System*; DSS) para problemas de seleção de portfólio no contexto de decisão multicritério, incorporando técnicas de aprendizagem de máquina e de algoritmos genéticos, para determinar e otimizar as constantes de escala (pesos) dos critérios.

1.3.2. Objetivos Específicos

- ✓ Fazer o levantamento bibliográfico de estudos publicados sobre os temas;
- ✓ Esquematizar um *framework* que descreva o modelo de análise (ou sistema de apoio à decisão) proposto;
- ✓ Detalhar cada fase do *framework*, especificando os procedimentos e os métodos/técnicas utilizados em cada etapa;
- ✓ Implementar computacionalmente a Fase de Seleção e Otimização, desenvolvendo um algoritmo genético e adaptando-o ao *software* já existente para seleção de portfólio;
- ✓ Desenvolver uma interface gráfica para usuários (especialistas) da fase implementada;
- ✓ Realizar testes de análise com dados de problemas reais.

1.4. RELEVÂNCIA E ORIGINALIDADE

Esta pesquisa tem sua relevância nas justificativas inerentes às contribuições que a mesma poderá proporcionar para os decisores e as organizações na resolução de problemas complexos. Doumpos e Zopounidis (2013) afirmam que a integração de MCDM com outros

paradigmas de Inteligência Artificial (IA), que incluem os métodos de ML e AG, é uma área de pesquisa interessante que fornece uma ampla gama de novos potenciais para melhorar a tomada de decisão.

Sendo assim, ressalta-se a necessidade de modelos e ferramentas mais sofisticadas para a resolução de problemas cada vez mais complexos. O processo de decisão depende do escopo e contexto do problema que, por vezes, requer um nível de sofisticação e confiabilidade além do que a capacidade de conhecimento humano pode, adequadamente, fornecer. Segundo Dean (2014) “devido à complexidade e tamanho da maioria dos problemas modernos, a melhor prática é tentar uma série de diferentes técnicas de modelagem ou algoritmos, usando diferentes configurações ou parâmetros”. Ou seja, é fundamental que os modelos tenham a flexibilidade para a integração dos métodos de diversas áreas, eliminando assim a lacuna existente no uso restrito de algoritmos de cada área para modelagem.

O modelo proposto surge como um processo alternativo ao habitual e tradicional processo de decisão multicritério, em que são levantados os dados iniciais (critérios; alternativas; desempenho de cada alternativa em cada critério), elicitado com o decisor as constantes de escala (pesos) e demais parâmetros, quando necessários, e aplicado um método multicritério, de acordo com a problemática requerida, para obtenção do resultado.

Nesse caso, o resultado é totalmente dependente das preferências do decisor, pois é ele quem define a ordem e as constantes de escala (pesos) dos critérios considerados no problema. Com o modelo proposto, a necessidade do decisor em definir a ordem e as constantes de escala (pesos) dos critérios é auxiliada, ou até mesmo substituída, pelo sistema, o que reduz a relação de dependência entre o decisor conseguir expressar suas preferências, para definir a ordem e o valor das constantes de escala (pesos) dos critérios, e assim obter resultados satisfatórios. Ou seja, o modelo de análise proposto torna o processo de decisão multicritério menos subjetivo.

Além disso, o modelo proposto considera a utilização das informações de experiências passadas, que são importantes fontes de conhecimento e aprendizagem, de uma forma eficiente para o processo de tomada de decisões. Spancher (2016) afirma que aprender a partir de exemplos anteriores, e com dados de treinamento adequados, a fim de decidir qual informação é mais confiável, é uma das capacidades que a tecnologia em aprendizagem de máquina pode oferecer, permitindo assim que outras abordagens, tal como a análise de decisão multicritério, apliquem um julgamento adequado em cada tipo de decisão, e o melhore ao longo do tempo. Desta forma, enfatiza-se que o modelo proposto nesta pesquisa

torna o processo de decisão menos subjetivo, uma vez que considera o aprendizado obtido com as experiências passadas, ao invés de levar em conta somente a opinião do decisor.

1.5. COMPLEXIDADE E LIMITAÇÕES DA PESQUISA

A complexidade da pesquisa consiste principalmente na forma como interligar e integrar os métodos de abordagens distintas, sendo que cada um possui suas fundamentações em teoremas e princípios já consolidados.

Com a base conceitual já definida em cada método, surge a dificuldade em conciliar as terminologias distintas, que na verdade significam a mesma coisa, e vice-versa, coisas diferentes que possuem a mesma terminologia. Por exemplo, em MCDM, o conceito de “critério” refere-se a “atributo”, termo usado em ML. O mesmo vale para os conceitos de “alternativas” e “instâncias”. Já o termo “normalização” é usado na análise de decisão multicritério, mas com significado diferente do que se entende por normalização na estatística.

Além da complexidade terminológica, a integração das abordagens em fases proporcionou um modelo flexível e dinâmico, podendo ser aplicado em diversas situações e com diferentes métodos, porém essa vantagem também trouxe como aspectos negativos, a complexidade de entendimento dos diversos caminhos a serem seguidos na análise.

Embora o modelo seja flexível, do ponto de vista da aplicabilidade em função das informações disponíveis, o mesmo é restritivo quanto ao tipo de problema a ser resolvido, ou seja, o modelo é limitado a resolver os problemas de seleção de portfólio, quando considerado vários critérios.

1.6. CLASSIFICAÇÃO DA PESQUISA

A pesquisa classifica-se quanto à natureza como uma pesquisa aplicada, pois busca gerar conhecimentos para aplicações práticas, dirigidos à solução de problemas específicos. Com relação ao ponto de vista da forma de abordagem ao problema, trata-se de pesquisa quantitativa, que considera a tradução de opiniões e números em informações, as quais serão classificadas e analisadas (Gil, 2008).

Com relação ao tipo de pesquisa do ponto de vista aos objetivos científicos, o estudo pode ser considerado como exploratória, pois envolve levantamento bibliográfico para entendimento e familiaridade do problema, e também como explicativo, uma vez que busca identificar fatores para elaboração de modelos. E do ponto de vista dos procedimentos técnicos, a pesquisa assume as formas bibliográfica, documental e experimental (Gil, 2008). Esta classificação geral da pesquisa é esquematizada na Figura 1.1.

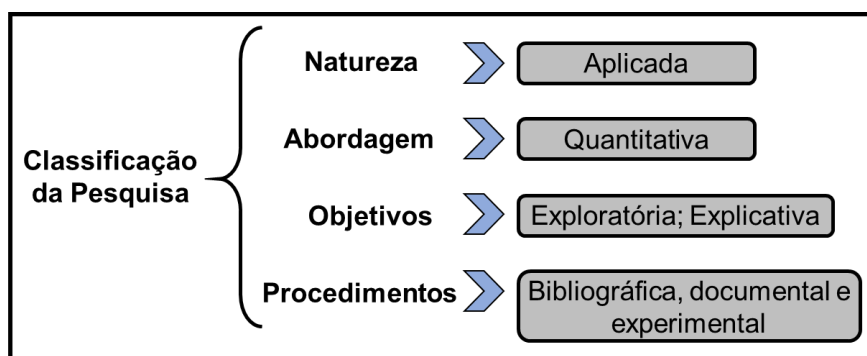


Figura 1.1 - Classificação da Pesquisa

1.7. ESTRUTURA DO DOCUMENTO

O trabalho está estruturado em sete capítulos: O Capítulo 1, a Introdução, que apresenta as motivações e justificativas para o desenvolvimento do trabalho, os objetivos do estudo, bem como também suas contribuições e limitações.

Os Capítulos 2 e 3 apresentam o Referencial Teórico e a Revisão de Literatura. No primeiro é definido o processo de modelagem do problema de seleção de portfólio, considerando as abordagens de decisão multicritério, de aprendizagem de máquina e dos algoritmos genéticos. O segundo expõe os trabalhos correlatos, com a revisão de literatura realizada nas principais bases de dados, organizando os artigos com uma análise bibliométrica e com a síntese do estado da arte.

Em seguida, o Capítulo 5 descreve a proposta do modelo em si, com a apresentação do framework para a análise de decisão multicritério em problemas de seleção de portfólio. Na sequência, os Capítulos 6 e 7 apresentam a implementação e a aplicação do modelo, respectivamente. No Capítulo 8 estão as considerações finais e propostas de futuras pesquisas, e logo após têm-se as referências, seguida dos Anexos.

2. REFERENCIAL TEÓRICO

Neste capítulo são abordadas algumas definições sobre o processo de modelagem, bem como classificadas e descritas algumas das principais técnicas e/ou métodos relacionados à análise multicritério, aprendizagem de máquina, e computação evolucionária.

Na seção sobre a análise multicritério, será definido o modelo aditivo para a problemática de portfólio, e também descrito o procedimento para determinação de pesos por *swing*, utilizado no método SMARTS (*Simple Multi-Attribute Rating Technique*), além de serem mencionados outros métodos possíveis para a parametrização do modelo aditivo.

Na seção sobre aprendizagem de máquina, serão apresentados e descritos os principais algoritmos classificadores (Redes Bayesianas, RB; Redes Neurais Artificiais, RNA; Árvore de Decisão, AD; Regressão Logística, RL; Máquina de Vetores de Suporte, MVS) para o reconhecimento de padrões na aprendizagem supervisionada, além das principais métricas de avaliação utilizadas para medir a eficiência de tais algoritmos, tais como a acurácia, precisão, *recall*, *F-score*, e área sob a curva ROC (*Receiver Operating Characteristic*). As técnicas para seleção dos atributos, que fornecem um *ranking* dos atributos, serão brevemente discutidas nesta seção também.

E, por fim, na seção sobre computação evolucionária, será apresentada a base teórica sobre os conceitos e definições utilizadas no algoritmo genético, tais como a definição das: a) representações de cromossomo e população; b) função de seleção; c) operações genéticas de crossover e mutação; d) critérios de avaliação e finalização.

2.1. ANÁLISE DE DECISÃO MULTICRITÉRIO PARA SELEÇÃO DE PORTFÓLIO

A Análise de Decisão Multicritério (*Multicriteria Decision Analysis*, MCDA) consiste em um conjunto de ferramentas que ajudam o decisor a organizar e sintetizar as informações necessárias e, conseqüentemente, sentir-se satisfeito, confortável e confiante com todos os critérios considerados, geralmente conflitantes entre si e, também, com o resultado da decisão, minimizando o potencial de “arrependimento”.

No contexto de MCDA existem os tipos de problemáticas a serem considerados em um processo de decisão. Cada tipo de problemática remete a um determinado objetivo que se pretende alcançar e, também, a determinados métodos que podem ser empregados para tanto.

Destacam-se as problemáticas, descritas e detalhadas por Almeida (2013), para a seleção, ordenação e classificação de alternativas, bem como para a de seleção de portfólio.

O problema de seleção de portfólio consiste na escolha de um ou mais projetos (alternativas, a_i) a serem executados dentre todos os n projetos (alternativas) disponíveis, ou seja, na decisão de qual(is) projeto(s) irá(ão) compor o portfólio, de tal forma a maximizar o retorno esperado (valor do portfólio, V), e sujeito a uma restrição, geralmente orçamentária, que limita a realização de todos os projetos (Belton e Stewart, 2002; Vetschera e Almeida, 2012). A função objetivo e a restrição desse problema podem ser descritas por (2.1) e (2.2), respectivamente, como segue:

$$\max V = \sum_{i=1}^n x_i v(a_i) \quad (2.1)$$

$$s. a: \sum_{i=1}^n x_i b_i \leq B \quad (2.2)$$

$$x_i \text{ binários} \quad (2.3)$$

onde x_i é uma variável binária (2.3), que indica se o projeto i será incluído ($x_i = 1$) ou não ($x_i = 0$) no portfólio; $v(a_i)$ é o valor do projeto i ; b_i representa o custo para se desenvolver o projeto i ; e B o orçamento total disponível (Kleinmuntz, 2007; Clemen e Smith, 2009). Desta forma, na função objetivo em (2.1) tem-se a maximização do valor do portfólio, que é dado pelo somatório do produto de x_i e $v(a_i)$, ou seja, o resultado da máxima soma dos valores dos projetos que são incluídos no portfólio. A restrição em (2.2) limita a inclusão de todos os projetos no portfólio, uma vez que o somatório dos custos, b_i , dos projetos incluídos deve ser menor ou igual ao orçamento total.

No contexto de decisão multicritério, esse problema é analisado considerando-se m critérios (c_j) e, para a obtenção da função de valor do projeto, $v(a_i)$, por meio de uma avaliação multicritério com método de agregação aditivo determinístico, é necessária a definição das constantes de escala (ou pesos, k_j) para cada um dos m critérios. O resultado dessa avaliação, conforme Almeida, *et al.* (2014) é da forma em (2.4).

$$v(a_i) = \sum_{j=1}^m k_j v_j(d_{ij}) \quad (2.4)$$

onde d_{ij} é o desempenho do projeto i em relação ao critério j ; v_j é a função valor marginal do critério j ; e k_j é a constante de escala (peso) do critério j , cuja soma deve ser igual a 1. Sendo assim, a função de valor do projeto descrita em (2.4) é obtida pelo somatório dos produtos entre os valores de constante de escala e o desempenho de cada projeto, em um determinado critério.

Um dos procedimentos para a definição das constantes de escala (pesos) dos critérios é a elicitaco direta com o decisor, com o procedimento de *trade-off* para a agregao aditiva da funo de valor (Almeida, 2013), tendo-se assim parâmetros que representam fielmente as preferências do decisor com relao aos critérios considerados para a análise. Porém, realizar esse procedimento com o decisor, na prática, muitas vezes é inviável, pois o decisor pode não estar disponível, ou então não trazer resultados consistentes, quando o decisor não tem definida de forma clara e objetiva a sua estrutura de preferência.

Outro procedimento para efetuar a agregao aditiva com a elicitaco das constantes de escala feita de forma mais simples e, portanto, com menor viés de erros de elicitaco, é o procedimento de pesos por *swing*, utilizado pelo método SMARTS. Este procedimento, segundo Almeida (2013) é conduzido pelas seguintes etapas:

- a) Ordenao dos critérios: consiste em perguntar ao decisor “*Suponha que voc seja obrigado a escolher uma alternativa que tem o pior desempenho em todos os critérios, e voc pudesse melhorar o desempenho desta alternativa em apenas um dos critérios para o valor máximo, qual critério voc escolheria?*”. O critério escolhido terá o maior valor de constante de escala, e o processo é repetido continuamente até que todos os critérios sejam avaliados;
- b) Obteno dos pesos: a partir da informao de ordem de importânca dos critérios, tem-se a obteno dos pesos, que consiste em perguntar ao decisor “*Supondo que o peso 100 [ou 10 ou 1, a depender da escala utilizada] é atribuído ao critério de maior peso, qual seria o peso do segundo critério mais importante?*”. Novamente, o processo é continuado até que todos os critérios tenham sido avaliados, e ao final é feita a normalizao dos pesos (diviso pela soma), de modo que a soma dos pesos seja igual a 1.

2.2 APRENDIZAGEM DE MÁQUINA SUPERVISIONADA

A aprendizagem de máquina é um ramo da IA, que engloba várias técnicas, algoritmos e teorias, tais como álgebra linear, probabilidade, grafos, entre outras. Estas visam um objetivo comum, o de possibilitar que uma máquina possa, através de observações de dados reais do modelo, reconhecer padrões, tomar decisões, classificar, prever acontecimentos, extrair e representar conhecimento.

Por exemplo, um servidor de *e-mails* pode aprender a classificar se um *e-mail* é ou não um *spam*, sendo que o conceito de *spam* à primeira vista é subjetivo, porém uma máquina pode aprender este conceito apenas com observações de humanos.

Segundo Alpaydin (2010), a aprendizagem de máquina é programar computadores para otimizar os parâmetros de um modelo usando uma base de treinamento ou experiências passadas, sendo que tal modelo é dito preditivo quando faz previsões sobre o futuro, ou descritivo quando extrai conhecimento dos dados, ou ambos. Para Russell e Norvig (2004), a ideia da aprendizagem de máquina não é ser um modelo estático, mas sim que a máquina aprenda continuamente, para que ela esteja sempre diminuindo os erros da tarefa designada (predição, classificação, etc.).

Alguns dos principais algoritmos utilizados para a tarefa de classificação, também denominados de algoritmos classificadores, são: Redes Bayesianas, RB; Redes Neurais Artificiais, RNA; Árvore de Decisão, AD; Regressão Logística, RL; Máquina de Vetores de Suporte, MVS. A seguir, será descrito cada um desses algoritmos.

2.2.1 Algoritmos classificadores: Redes Bayesianas

Uma rede bayesiana é um modelo gráfico que consiste de duas partes, G e P , onde G é um gráfico acíclico dirigido (*Directed Acyclic Graph*; DAG), cujos nós representam variáveis aleatórias e os arcos entre nós representam dependência condicional das variáveis aleatórias; e P é um conjunto de distribuições das probabilidades condicionais (*Conditional Probability Distribution*; CPD) para cada nó condicionada em seus pais (Dean, 2014).

As distribuições de probabilidade condicional para cada nó podem ser obtidas de maneira prévia, a partir do conhecimento de especialistas, ou então aprendidas, a partir de dados. Koller e Friedman (2009) ressaltam que o gráfico G pode ser visto de duas maneiras diferentes, porém equivalentes: “como uma estrutura de dados que fornece o esqueleto para

representar uma distribuição conjunta de forma compacta numa forma factorizada; como uma representação compacta para um conjunto de suposições de independência condicional sobre uma distribuição”.

A Figura 2.1 mostra um exemplo de uma rede Bayesiana para o problema de compra de roupa, que leva em consideração os seguintes atributos: a estação do ano; a localização; o peso e a cor do tecido. A resolução deste problema fornece, por exemplo, as probabilidades de comprar ou um casaco, ou uma bermuda, ou uma camisa, levando em consideração a probabilidade dos atributos anteriormente citados. Sendo assim, se a estação do ano for primavera, e a localização for New York, a probabilidade de comprar uma bermuda será maior do que a probabilidade de comprar um casaco.

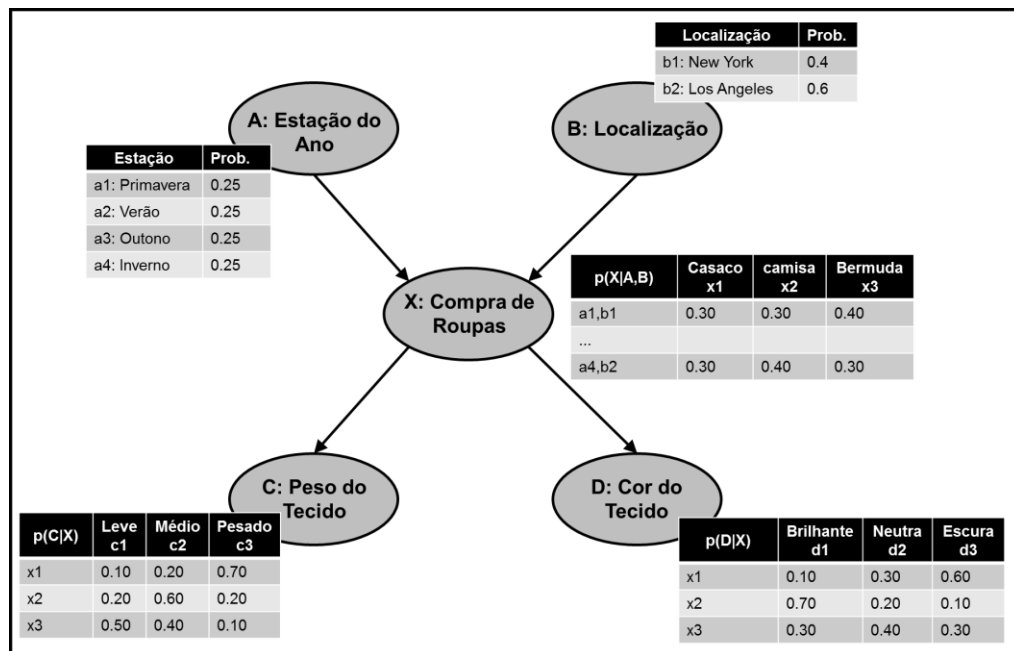


Figura 2.1 - Exemplo de uma rede bayesiana aplicada ao problema de compra de roupa

Apesar do nome, as redes bayesianas não implicam necessariamente que elas dependam de estatísticas bayesianas. Pelo contrário, elas são assim chamadas porque usam a regra de Bayes para inferência probabilística. É possível usar estatísticas bayesianas para aprender uma rede Bayesiana, mas há também muitas outras técnicas que estão mais intimamente relacionadas com os métodos estatísticos tradicionais (Dean, 2014). Russell *et al.* (1995) sugerem o método do gradiente descendente para a aprendizagem em redes bayesianas, análogo ao caso de aprendizagem de redes neurais, que pode ser aplicado para encontrar o conjunto ideal de pesos (valores de probabilidade).

2.2.2 Algoritmos classificadores: Redes Neurais Artificiais

Outra técnica muito utilizada e estudada em aprendizado de máquina são as Redes Neurais Artificiais, que foi inspirada no funcionamento do cérebro humano em que neurônios se conectam e passam informações entre si. Essa técnica pode ser usada tanto para resolver problemas de classificação, que possuem classes como saídas, quanto para problemas de regressão, que possuem valores de predição como saídas (Russell e Norvig, 2004).

Mais especificamente, uma rede neural artificial é um modelo matemático construído a partir de uma rede de nós, que representam funções de ativação, e conexões, que representam os fluxos por onde as variáveis passam. Cada nó da rede, também chamado de neurônio, é responsável por transformar as informações que entram nele e produzir um resultado como saída através de uma função matemática, portanto os neurônios têm como entradas várias variáveis e como saída somente um valor real. Cada conexão representa qual variável entra em qual neurônio.

Uma das arquiteturas mais utilizadas de redes neurais é a "*multilayer perceptron*" (Dean, 2014), onde existem diferentes camadas que se conectam entre si, entretanto os neurônios da mesma camada não se conectam uns com os outros. A Figura 2.2 mostra uma rede "*multilayer perceptron*" com uma camada de entrada, duas camadas internas/escondidas (*hidden layer*) e uma camada de saída.

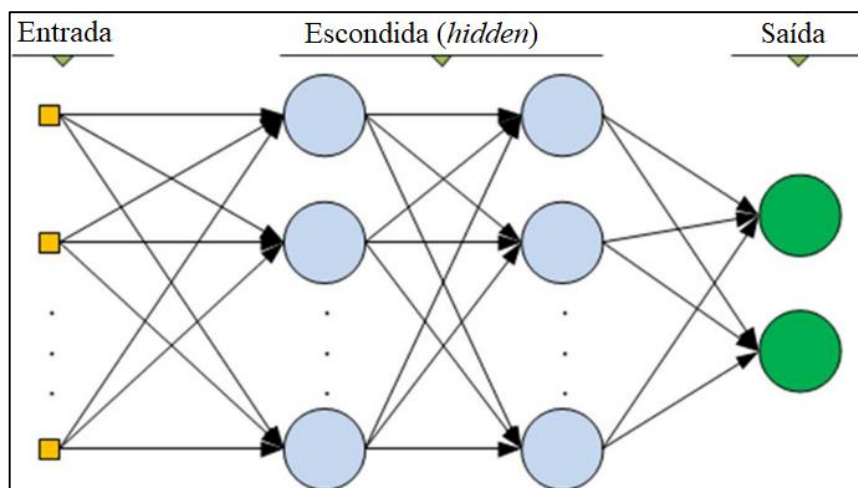


Figura 2.2 - Rede Neural Artificial, *Multilayer Perceptron*

Fonte da Imagem: <https://www.hackerearth.com/blog/wp-content/uploads/2017/01/multi-1024x533.jpg>

Na camada de entrada temos as variáveis de entrada do problema, já na camada de saída temos o resultado final tanto para classificação quanto para regressão. Note que é possível haver mais de uma variável na camada de saída, ou seja, várias saídas.

Os procedimentos e algoritmos de treinamento são responsáveis por ajustar os parâmetros das funções de cada neurônio de forma a minimizar os erros de saída da rede e, para isso, são utilizados dados de treinamento em que as entradas e saídas corretas são conhecidas. Um algoritmo de treinamento bastante utilizado em redes neurais é o "*Backpropagation*", o conceito principal desse algoritmo é medir qual é efeito dos pesos entre os neurônios das camadas e, para isso, o erro da saída é retro propagado para as camadas anteriores possibilitando medir esse efeito e ajustar os pesos para minimizar o erro de saída (Hastie, Tibshirani e Friedman, 2008).

2.2.3 Algoritmos classificadores: Árvores de Decisão

As árvores de decisão são modelos populares em aprendizado de máquina que descrevem algoritmos de decisão em forma de uma árvore de possibilidades de valores das entradas. Conforme as entradas mudam de valores, os caminhos da árvore se alteram, e a saída final também se altera com o objetivo de capturar uma estratégia ótima de mapear entradas em saídas (Dean, 2014; Larose e Larose, 2015).

A árvore de decisão possui três elementos: as variáveis, "nós" de decisão; os "nós" finais, também chamados de folhas; e as conexões. A Figura 2.3 mostra um exemplo simples de como usar a árvore de decisão para prever se um cliente irá comprar ou não um carro, considerando as variáveis "idade"; "estudante" e "taxa de crédito". Pelo exemplo, se for um cliente de meia idade (idoso), a compra do carro será feita somente se a taxa de crédito for excelente.

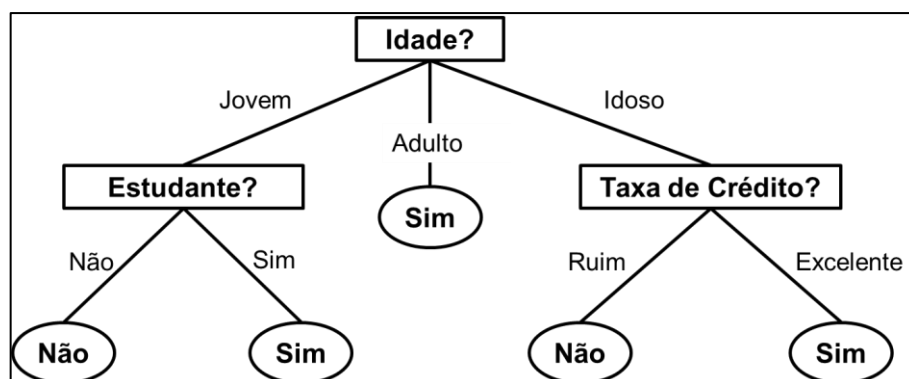


Figura 2.3 - Árvore de Decisão, exemplo para compra de carro
 Fonte: <https://zhengtianyu.files.wordpress.com/2013/12/decision-tree.png>

As variáveis são parte do modelo e seus valores são as entradas do sistema, os “nós” de decisão implementam regras de decisão sobre uma variável podendo mudar o fluxo da árvore de acordo com essa regra, as conexões definem quais são os possíveis caminhos entre as regras de decisão, e as folhas são o resultado final da árvore (Zaki e Meira Junior, 2014).

As saídas das árvores de decisão podem ser tanto classes, chamadas de árvores de classificação, englobando problemas de classificação e regressão, quanto números reais, chamadas de árvores de regressão, englobando problemas de regressão. O treinamento das árvores tem por objetivo criar regras nos nós de decisão que mais se ajustam ao cenário proposto. Os algoritmos mais utilizados para aprender esse tipo de modelo são os algoritmos gulosos baseados em particionamento recursivo (Sugiyama, 2015).

Com o objetivo de melhorar as estimativas, Dean (2014) afirma que, em muitas ocasiões, várias árvores de decisão são utilizadas em conjunto para melhorar as estimativas das árvores. O resultado final, portanto, considera os resultados individuais de cada árvore assim o problema comum de "overfit" dos classificadores e estimadores singulares.

2.2.4 Algoritmos classificadores: Regressão Logística

A regressão logística é uma técnica utilizada em aprendizagem de máquina supervisionado para abordar o problema classificação. Nesta técnica o objetivo é usar a função logística com seus respectivos parâmetros (β ; Equação 2.5) para inferir se uma instância pertence ou não a uma determinada classe. Como essa é uma técnica de aprendizagem supervisionada, a partir de um conjunto de dados, no qual cada instância é marcada com a classe verdadeira, é possível ajustar os parâmetros da função logística ao ponto de maximizar o número de classificações corretas (Sugiyama, 2015).

$$p_i = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_{1,i} + \dots + \beta_k x_{k,i})}} \quad (2.5)$$

Nota-se que na equação, a entrada é um vetor multidimensional, portanto a regressão logística engloba cenários de aprendizagem com várias variáveis de entrada. Analogamente, se a entrada é multivariada, a saída da função logística é univariada, portanto uma única função logística informa a probabilidade de uma determinada instância pertencer ou não a uma classe. Em um cenário com mais de uma classe é possível aliar a regressão logística a

técnicas como "*one-against-all*" em que para cada classe é usada uma função logística diferente (Hastie, Tibshirani e Friedman, 2008).

Por não possuir solução ótima analítica, os ajustes de pesos da função logística para os dados de treinamento são feitos geralmente por algoritmos iterativos como o "*Gradient Descent*" e o "*Newton's Method*".

2.2.5 Algoritmos classificadores: Máquina de Vetores de Suporte

Em aprendizado de máquina, máquinas de vetores de suporte (*Support Vector Machines, SVM*) são modelos de classificação que ajustam hiperplanos em um espaço n -dimensional separando o conjunto de dados em duas classes, como mostra a Figura 2.4. O ajuste é feito de modo que a distância entre hiperplano e o ponto mais próximo a ele no conjunto de dados seja a maior possível (Zaki e Meira Junior, 2014).

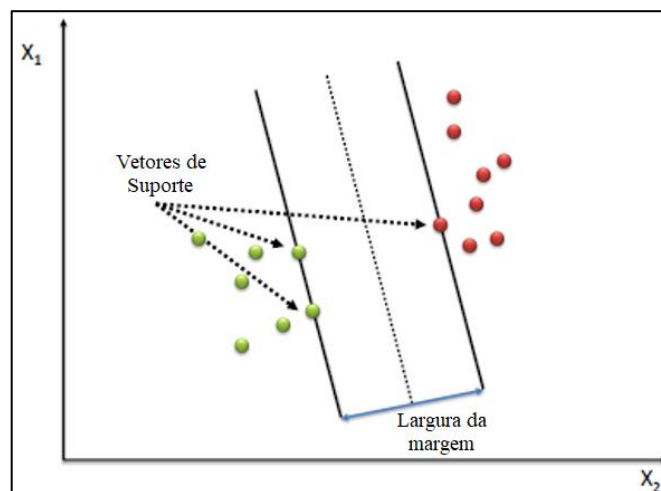


Figura 2.4 - Esquema da Máquina de Vetores de Suporte (*Support Vector Machine, SVM*)

Fonte: https://www.saedsayad.com/images/SVM_2.png

Como o classificador da SVM é linear, um mapeamento do espaço n -dimensional para um espaço de dimensão maior que " n " pode ser feito para abordar problemas não lineares, de forma que neste novo espaço dimensional o conjunto de dados seja linearmente separável. Essa transformação não-linear é em geral computacionalmente custosa por isso são feitas com as funções "*kernel*" (referência), essas funções evitam que seja necessário mapear todo o conjunto de dados para um novo espaço dimensional. Após a transformação, aplica-se o mesmo classificador linear para o novo espaço com maiores dimensões (Dean, 2014).

2.2.6 Métricas de avaliação dos algoritmos classificadores

Algoritmos de aprendizado de máquina podem ter seu desempenho avaliado por diversos motivos como, comparar algoritmos, verificar a evolução de um algoritmo, ajustar melhores parâmetros e etc. Para a avaliação de algoritmos de classificação, as seguintes métricas são as mais utilizadas: Acurácia; Precisão; Recall; F-Score; "Area Under Curve". Os cálculos dessas métricas são feitos com base na matriz de confusão (Quadro 2.1), que ilustra os acertos e erros do classificador (Freitas, 2003).

		Classe Prevista	
		A	B
Classe Real	A	VP	FN
	B	FP	VN

Quadro 2.1 - Matriz de Confusão (2x2)

Partindo-se dos valores da matriz de confusão (Verdadeiro Positivo, VP; Verdadeiro Negativo, VN; Falso Positivo, FP; e Falso Negativo, FN), Freitas (2003) define as métricas de:

- a) Acurácia: Representa a porcentagem total de acertos do classificador, tanto positivos quanto negativos (Equação 2.6).

$$Acurácia = \frac{VP+VN}{VP+VN+FP+FN} \quad (2.6)$$

- b) Precisão: Representa a porcentagem de acertos positivos, com relação a todas as instâncias classificadas como positivas (Equação 2.7).

$$Precisão = \frac{VP}{VP+FP} \quad (2.7)$$

- c) Sensibilidade ou *Recall*: Representa a porcentagem de acertos positivos, com relação a todas as instâncias positivas de fato (Equação 2.8).

$$Recall = \frac{VP}{VP+FN} \quad (2.8)$$

- d) Especificidade: é a proporção de verdadeiros negativos, ou seja, a capacidade do modelo em prever corretamente a ausência da condição para casos que realmente não a têm (Equação 2.9)

$$\text{Especificidade} = \frac{VN}{VN+FP} \quad (2.9)$$

- e) *F-Score*: é definido como a média harmônica da Precisão e do *Recall*, unindo as duas métricas em uma só (Equação 2.10). Um algoritmo com Precisão alta e *Recall* baixo tem robustez, porém perde uma parcela de instâncias que são difíceis de classificar, portanto um algoritmo com *F-Score* alto tende a ter um balanço entre precisão e *recall*.

$$F - Score = \frac{2 \cdot recall \cdot precisão}{recall + precisão} \quad (2.10)$$

- f) Área sob a curva: Mais conhecido como AUC (*Area Under Curve*), essa métrica se vale do gráfico *Receiver operating characteristic* (ROC), exemplificado na Figura 2.5, que traça uma curva nos eixos "Taxa de Verdadeiros Positivos" *versus* "Taxa de Falsos Positivos" para vários valores de limiares (*threshold*) do classificador. Quanto maior for a área abaixo da curva ROC melhor será o classificador.

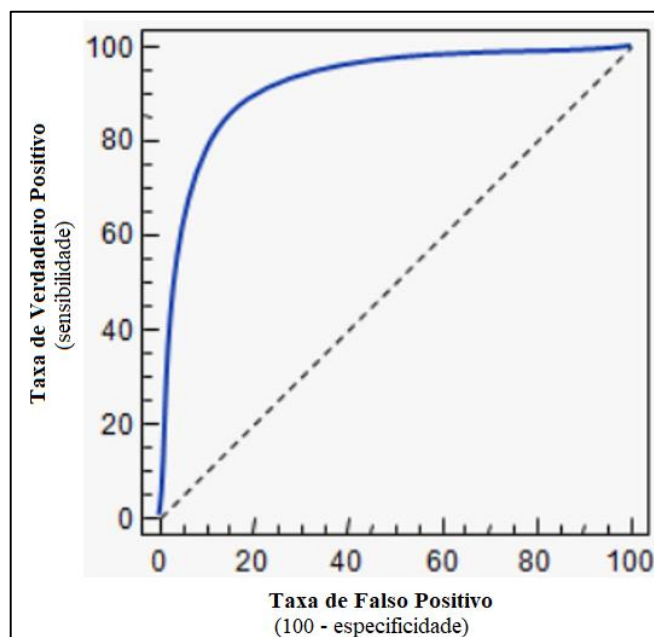


Figura 2.5 - Curva ROC (*Receiver Operating Characteristic*)

Fonte: <https://i.stack.imgur.com/5x3Xj.png>

2.2.7 Seleção de Atributos

Em problemas de aprendizado de máquina com muitos atributos, pode ser útil (e em alguns casos pode até melhorar o desempenho do algoritmo) usar a técnica de seleção de atributos. Essa técnica parte do princípio que nem todos os atributos são úteis para se obter uma saída com o melhor desempenho possível, e por vezes alguns atributos podem até atrapalhar o desempenho do algoritmo (Freitas, 2003).

Conforme Freitas (2003) existem algumas vantagens de usar esse tipo de técnica, como: simplificação dos modelos; menor tempo de treinamento; redução das dimensões poderá requerer menos instâncias de treinamento; redução do efeito de *overfit*.

Dessa forma, a Seleção de Atributos escolhe somente os atributos mais relevantes para o cenário proposto de duas principais formas, o método “*Wrapper*” e o “*Ranker*”. O método *Wrapper* utiliza o próprio algoritmo de aprendizado de máquina escolhido (regressão logística, redes neurais, entre outros) para testar possibilidades de treinamento com menos variáveis, sendo que ao final é escolhido o conjunto de atributos com o maior desempenho.

O método *Ranker* verifica métricas estatísticas como Informação Mútua, Informação Mútua Pontual e Correlação de Pearson, para cada par atributo/saída, de forma que se crie um *ranking* das variáveis que mais têm influência sobre a saída. Com esse *ranking* pode-se escolher somente um conjunto limitado dos primeiros atributos para criar o novo modelo (Larose e Larose, 2015).

2.3 COMPUTAÇÃO EVOLUCIONÁRIA

A computação evolucionária, ou computação evolutiva, é um ramo de pesquisa emergente da IA, que propõe um paradigma para solução de problemas inspirado na seleção natural. Ou seja, a computação evolucionária compreende diversos algoritmos inspirados no princípio Darwiniano da evolução das espécies e na genética. São algoritmos probabilísticos, que fornecem um mecanismo de busca paralela e adaptativa baseado no princípio de sobrevivência dos mais aptos e na reprodução (Larose e Larose, 2015).

Existem inúmeros problemas para os quais deseja-se desenvolver um algoritmo eficiente. Muitos destes, são problemas de otimização numérica ou combinatória, para os quais é frequentemente possível encontrar um algoritmo que ofereça uma solução ótima ou aproximadamente ótima. A computação evolucionária oferece algoritmos gerais, dentre os

quais se enquadram os Algoritmos Genéticos, a Programação Genética e a Eletrônica Evolucionária, que são aplicados em problemas complexos, com grandes espaços de busca, de difícil modelagem, ou para os quais não há um algoritmo eficiente disponível.

2.3.1 Algoritmo Genético

Os Algoritmos Genéticos foram concebidos em 1960 por John Holland, com o objetivo inicial de estudar os fenômenos relacionados à adaptação das espécies e seleção natural, incorporando tais conceitos aos computadores, simulando processos naturais de sobrevivência e reprodução das populações.

Compreendem uma técnica na qual os cromossomos são palavras binárias que codificam soluções para um problema, e possuem uma larga aplicação em muitas áreas científicas, como otimização de soluções, aprendizado de máquina, desenvolvimento de estratégias, problemas de engenharia, análise de modelos econômicos, dentre outras aplicações.

Segundo Freitas (2003), para a implementação de um AG é necessário determinar:

- a) representações das possíveis soluções do problema, no formato de um código genético;
- b) a população inicial que contenha diversidade suficiente para permitir ao algoritmo combinar características e produzir novas soluções;
- c) existência de um método para medir a qualidade de uma solução potencial, denominado de *fitness*;
- d) um procedimento de combinação de soluções para gerar novos indivíduos na população, podendo este ser operações de *crossover*, mutação ou reparação;
- e) um critério de escolha das soluções que permanecerão na população, ou que serão excluídos desta, sendo que no AG generacional toda a população é substituída, enquanto que no AG *steady-state* é introduzido/excluído apenas um (ou dois) indivíduo(s) por iteração;
- f) um critério de parada do algoritmo, podendo esta ser pelo número de iterações, acréscimo na melhor solução, ou então até atingir um determinado valor de *fitness* desejado.

A execução do algoritmo pode ser resumida nas seguintes etapas básicas:

- Passo 1: Criar a População com o número de indivíduos desejado;

- Passo 2: Calcular o *fitness* (aptidão), podendo ser de maximização ou minimização;
- Passo 3: Colocar a população em ordem, segundo o fitness de cada indivíduo (do melhor para o pior);
- Passo 4: Utilizar a função de seleção (roleta; *ranking* ou função *select*, sendo esta última descrita pela Equação 2.11), para selecionar dois indivíduos da população que serão os “pais”;

$$Select(R) = \left\{ r_j, \text{tal que } j = m + 1 - \left\lceil \frac{-1 + \sqrt{1 + 4 \cdot RND(m^2 + m)}}{2} \right\rceil \right\} \quad (2.11)$$

onde r_j representa o j -ésimo cromossomo dentro da população de tamanho " m "; RND é um número randômico entre 0 e 1; e o símbolo $\lceil x \rceil$ significa a operação matemática de arredondamento para cima do valor x .

- Passo 5: Realizar as operações genéticas de cruzamento e mutação (Figura 2.6), para a criação dos filhos e, caso necessário, efetuar a reparação;

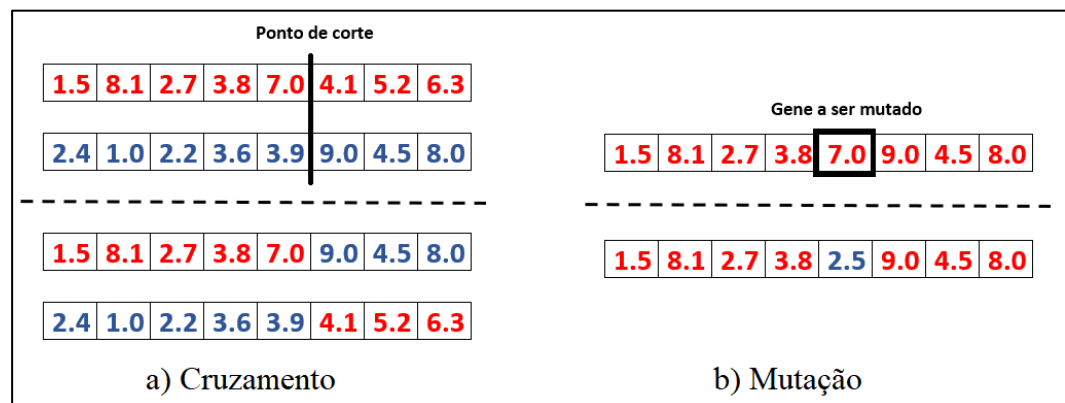


Figura 2.6 - Operações de Cruzamento e Mutação

- Passo 6: Gerar uma nova população inserindo os filhos na população obtida no Passo 3.
- Passo 7: Verificar o critério de parada, se não atendido voltar ao Passo 3, caso contrário finalizar o procedimento.

3. REVISÃO DE LITERATURA E TRABALHOS CORRELATOS

Este capítulo apresenta a pesquisa realizada com a finalidade de encontrar e revisar a literatura existente, em termos de artigos publicados, sobre os temas em foco neste trabalho. Para isso, foram seguidos os passos descritos na Figura 3.1, que exemplifica de forma resumida as etapas da revisão sistemática de literatura, composta da análise bibliométrica e da análise de conteúdo.

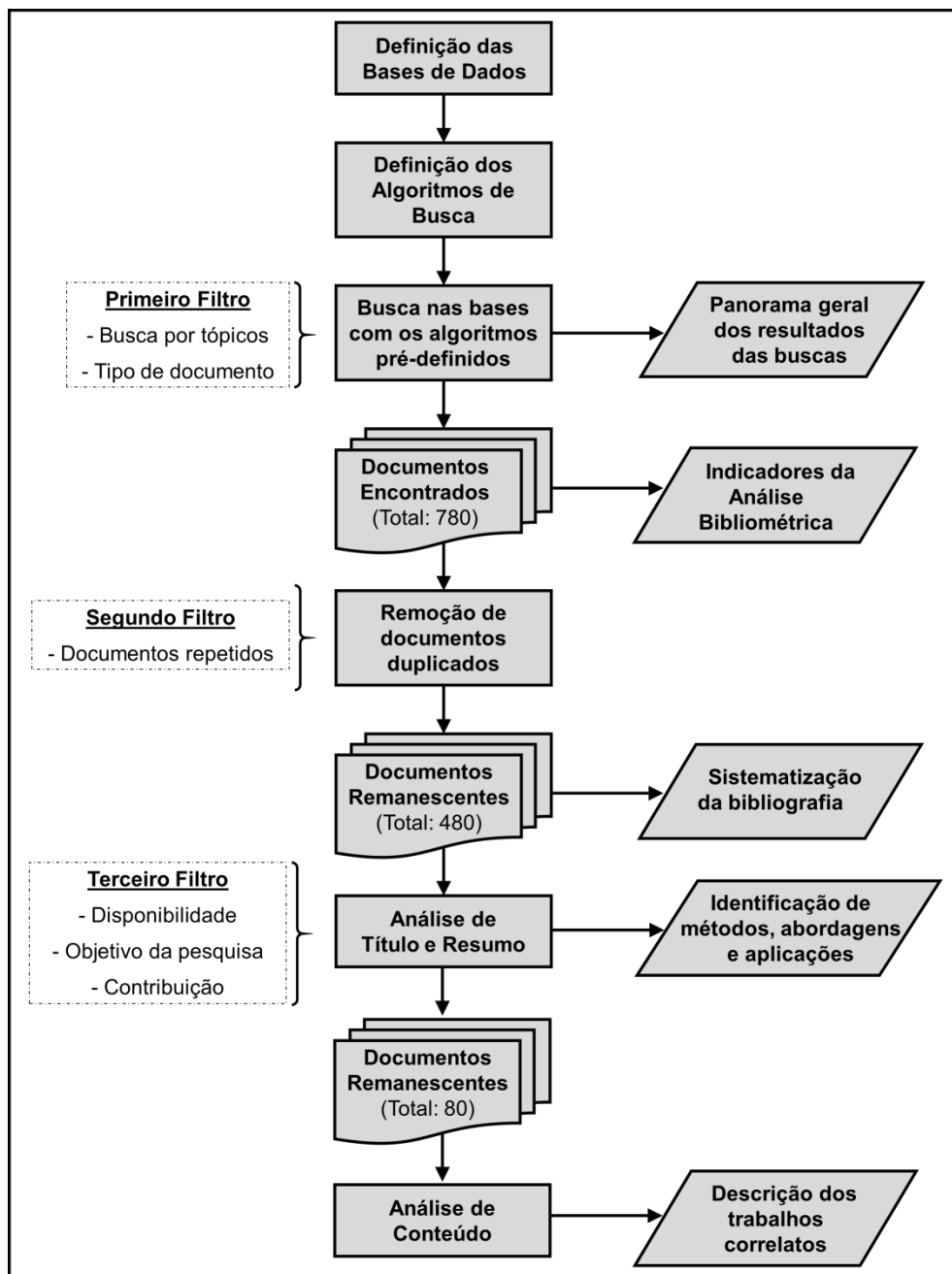


Figura 3.1 - Framework do processo de revisão sistemática, análise bibliométrica e análise de conteúdo

Como pode ser visto na Figura 3.1, a primeira etapa para a revisão sistemática de literatura consiste na definição das bases de dados. Foram consideradas como fonte de busca três das principais bases de dados de artigos indexados, que são a *Science Direct*, a *Scopus* e a Coleção Principal do *Web of Science*.

Os termos de busca foram definidos a partir de conceitos e terminologias referentes aos temas da pesquisa, de forma ampla, para abranger o máximo possível de artigos publicados, favorecendo assim uma revisão e uma análise completas e robustas. Sendo assim, os termos considerados, cada um relacionado a uma área diferente de conhecimento aplicado nesta pesquisa, foram:

- “*portfolio selection*”: que define o tipo de problema a ser resolvido;
- “*multicriteria decision*” (e termos similares: “*multi-criteria decision*”, e “*multi-attribute decision*”): que expõe a condição de se ter vários critérios/atributos na análise;
- “*genetic algorithm*”: que identifica a metaheurística utilizada;
- “*machine learning*” (e termo similar “*data mining*”): que identifica as técnicas utilizadas.

Como a inclusão dos vários termos em um único algoritmo de busca resulta em uma pesquisa mais restritiva, resultando em poucos artigos publicados, é necessário proceder com mais de uma busca, incluindo a combinação de termos nos algoritmos, para se ter uma pesquisa mais ampla.

E, de fato, a busca feita com os quatro termos em um único algoritmo [“*portfolio selection*” AND “*multicriteria decision*” (OR “*multi-criteria decision*” OR “*multi-attribute decision*”) AND “*genetic algorithm*” AND “*machine learning*” (OR “*data mining*”)], nas três bases de dados (*Science Direct*, *Scopus* e *Web Of Science*), não resultou em nenhum documento encontrado. E as combinações feitas com três dos quatro termos, em dois algoritmos de busca diferentes, [“*portfolio selection*” AND “*multicriteria decision*” (OR “*multi-criteria decision*” OR “*multi-attribute decision*”) AND “*genetic algorithm*”] e [“*portfolio selection*” AND “*multicriteria decision*” (OR “*multi-criteria decision*” OR “*multi-attribute decision*”) AND “*machine learning*” (OR “*data mining*”)], resultaram no total em apenas sete e um documento (s) encontrado (s), respectivamente.

Somente com essa quantidade de artigos não é possível fazer uma análise bibliométrica adequada, porém isso demonstra que são pouquíssimos os estudos similares ao proposto nesta pesquisa, evidenciando assim a contribuição por originalidade e ineditismo,

uma vez que nenhum destes trabalhos possui o mesmo objetivo e o mesmo uso conjunto das técnicas que esta pesquisa apresenta.

Então, para ter uma busca mais ampla e, conseqüentemente, uma análise bibliométrica mais adequada, foram consideradas as seguintes combinações dos termos para a composição dos algoritmos de busca:

a) [*“portfolio selection”* AND *“multicriteria decision”* (OR *“multi-criteria decision”* OR *“multi-attribute decision”*)]: para capturar os estudos cujos focos sejam o cerne desta pesquisa, ou seja, análise de decisão multicritério em problemas de seleção de portfólio;

b) [*“portfolio selection”* AND *“genetic algorithm”*]: para verificar os estudos que empregaram a otimização por algoritmos genéticos no contexto de seleção de portfólio;

c) [*“portfolio selection”* AND *“machine learning”* (OR *“data mining”*)]: para identificar os estudos que estão empregando conceitos de aprendizagem de máquina em problemas de seleção de portfólio;

d) [*“portfolio selection”* AND (*“multicriteria decision”* OR *“multi-criteria decision”* OR *“multi-attribute decision”* OR *“genetic algorithm”* OR *“machine learning”* OR *“data mining”*)]: para identificar de forma mais ampla os estudos que empregam decisão multicritério, ou algoritmos genéticos, ou então aprendizagem de máquina, no contexto de seleção de portfólio.

A busca nas bases de dados com estes algoritmos (terceira etapa do processo de revisão sistemática; Figura 3.1) foi feita considerando-se a busca por “tópicos”, ou seja, por documentos que apresentam os termos ou no título, ou no resumo ou nas palavras-chave, e também apenas por documentos do tipo *“article”*, *“conference paper”*, *“book chapter”* e *“review”*, sendo este o primeiro filtro do processo, em que excluiu-se da busca todos os documentos que não se enquadraram nestas condições. Como resultado parcial desta etapa, tem-se um panorama geral das buscas, representado pelas Figuras 3.2, que demonstra a quantidade de documentos encontrados por algoritmo e por base.

O resultado das buscas com a combinação dos termos em diferentes algoritmos, apresentado na Figura 3.2, evidencia o maior número de pesquisas realizadas com a aplicação de algoritmos genéticos em problemas de seleção de portfólio. Também, tem-se uma grande quantidade de trabalhos que consideram o contexto de multicritérios/multiobjetivo no referido tipo de problema, sendo superior as pesquisas que empregam técnicas de aprendizagem de máquina. Isso deve-se principalmente ao fato de que o desenvolvimento e a aplicação das

técnicas de aprendizagem de máquina são relativamente mais recentes, porém com um crescimento potencial na literatura nos últimos anos.

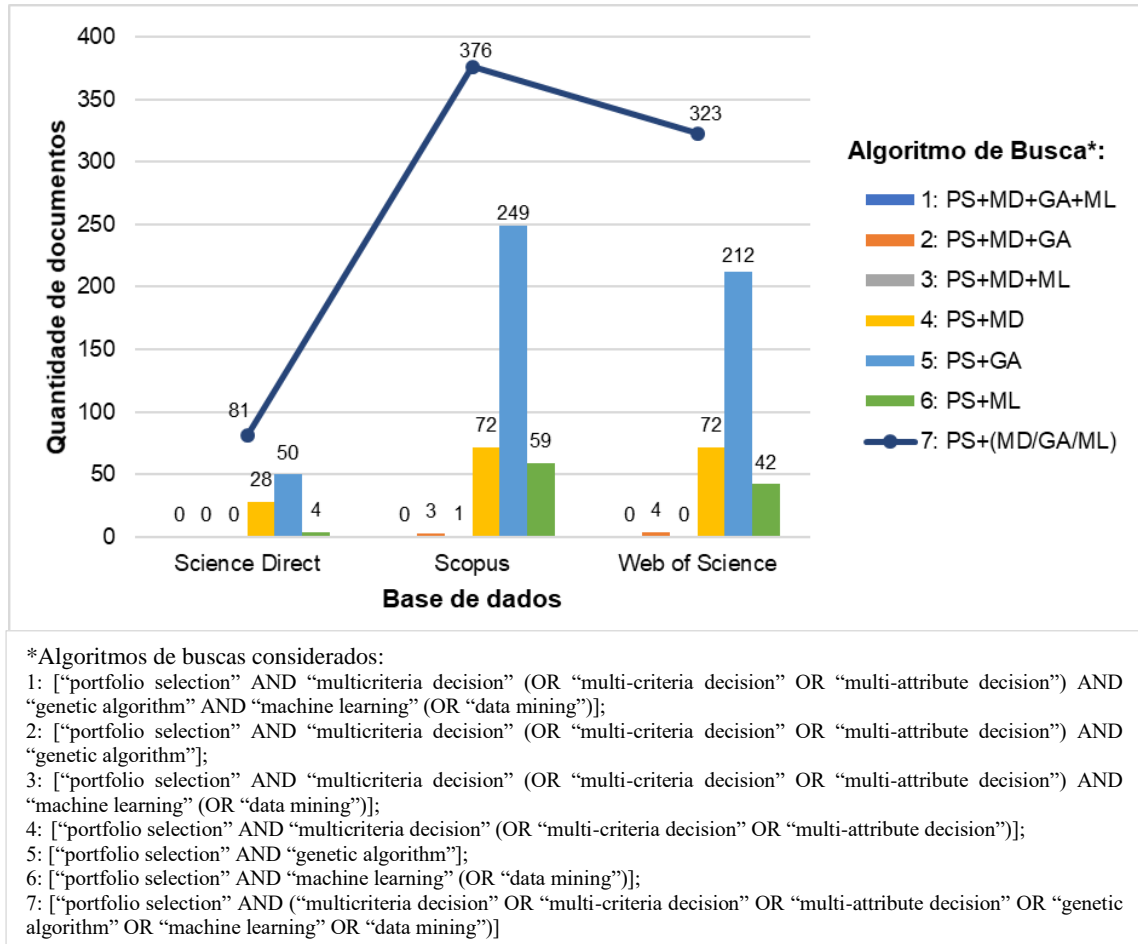


Figura 3.2 - Quantidade de documentos encontrados nas bases de dados com diferentes algoritmos de busca

Percebe-se também que as buscas nas bases de dados com o algoritmo 7: ["*portfolio selection*" AND ("multicriteria decision" OR "multi-criteria decision" OR "multi-attribute decision" OR "genetic algorithm" OR "machine learning" OR "data mining")] foram mais amplas, englobando praticamente todos os artigos contidos nas buscas com os demais algoritmos. Por conta disso, a continuidade do processo de revisão sistemática de literatura deu-se apenas com os documentos encontrados com este algoritmo de busca, que resultou em: 81 documentos na base *Science Direct*; 376 documentos na base *Scopus*; 323 documentos na base *Web of Science*; e um total de 780 documentos encontrados.

Desta forma, a análise bibliométrica, que considerou as duas bases com maior número de documentos (*Scopus* e *Web of Science*), resultou em indicadores do número de documentos por autor, ano, localização, área do conhecimento, e revista/periódico, também do número de

citações por documento e ano. Estes indicadores estão apresentados em forma de tabelas de gráficos na seção 3.1.

Todos os 780 documentos encontrados, provenientes das três bases de dados, foram exportados para o *software* gerenciador de bibliografias (EndNote), onde foi realizada a remoção de documentos duplicados, sendo este o segundo filtro do processo de revisão. Ao final desta etapa, restaram-se 480 documentos únicos, que foram sistematizados em uma planilha, e organizados quanto ao autor, título, ano de publicação, resumo e palavras-chave.

Com a leitura do título e resumo dos artigos, foi possível identificar os conceitos, as abordagens e os métodos utilizados, bem como aplicações realizadas pelos autores. E, a partir disso, foram elaborados gráficos e tabelas para resumir tais informações, sendo apresentados na seção 3.2. Foi também nesta etapa que ocorreu a aplicação do terceiro filtro, com a identificação dos trabalhos cujos objetivos e contribuições foram relevantes e/ou similares a esta pesquisa, e com a exclusão daqueles trabalhos cuja disponibilidade de acesso total ao conteúdo não foi permitida. Ao final, restaram-se 80 documentos para a análise de conteúdo.

A análise de conteúdo, última etapa do processo de revisão sistemática de literatura, foi feita com leitura completa dos artigos sendo, posteriormente, apresentadas as descrições de trabalhos correlatos mais importantes, na seção 3.3.

3.1. ANÁLISE BIBLIOMÉTRICA

Como já mencionado anteriormente, a análise bibliométrica foi realizada com o resultado da busca utilizando-se o algoritmo 7: [*“portfolio selection”* AND (*“multicriteria decision”* OR *“multi-criteria decision”* OR *“multi-attribute decision”* OR *“genetic algorithm”* OR *“machine learning”* OR *“data mining”*)], em duas bases de dados, *Scopus* e *Web of Science*, que tiveram um maior número de documentos encontrados, 376 e 323, respectivamente.

Essa quantidade total de documentos encontrados nas bases de dados teve uma variação no decorrer dos últimos 10 anos, como pode ser visto na Figura 3.3. Observa-se também uma tendência geral de crescimento na quantidade de documentos publicados, em ambas as bases de dados, sendo que o decréscimo em 2018 se deve, principalmente, ao fato de o ano não ter acabado no momento em que a busca foi realizada.

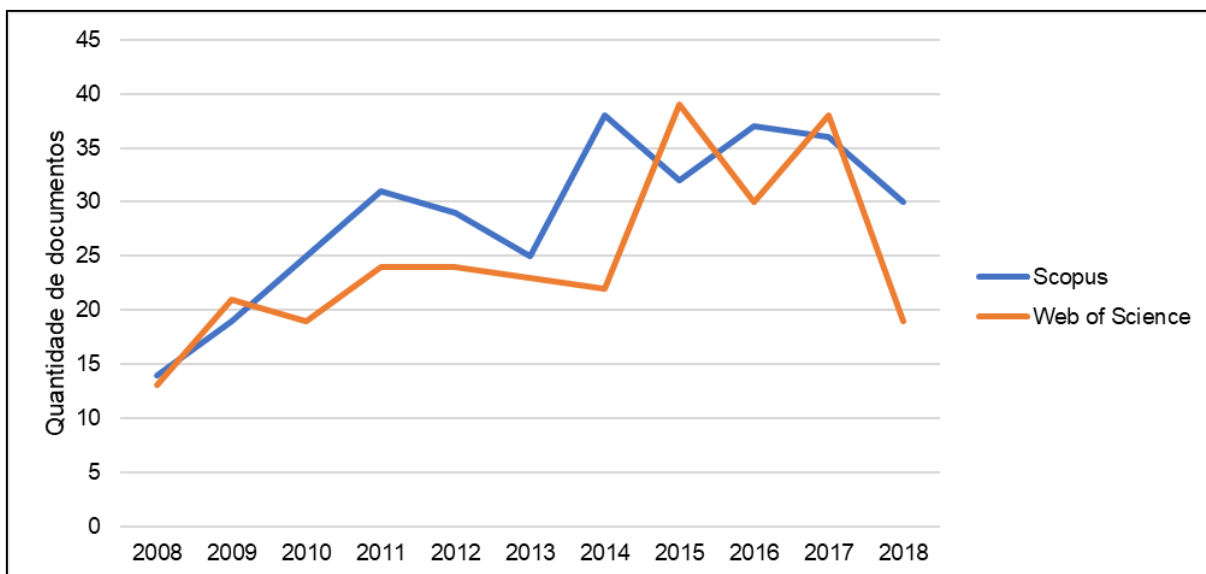


Figura 3.3 - Quantidade de documentos encontrados por ano nas bases de dados Scopus e Web of Science

Esta tendência de crescimento na quantidade de artigos publicados também é verificada no número de citações que tais documentos receberam durante os anos, como mostra a Figura 3.4. Em 2017 o número de citações dos trabalhos atingiu o auge, com aproximadamente 670 citações, em ambas as bases de dados. Tanto o crescimento da quantidade de publicações, quanto o aumento no número de citações dos documentos relacionados ao tema da pesquisa, demonstram uma evolução positiva para os estudos que envolvem os diversos métodos e técnicas para o problema de seleção de portfólio, ou seja, este é ainda um tema bastante procurado e difundido pelos pesquisadores.

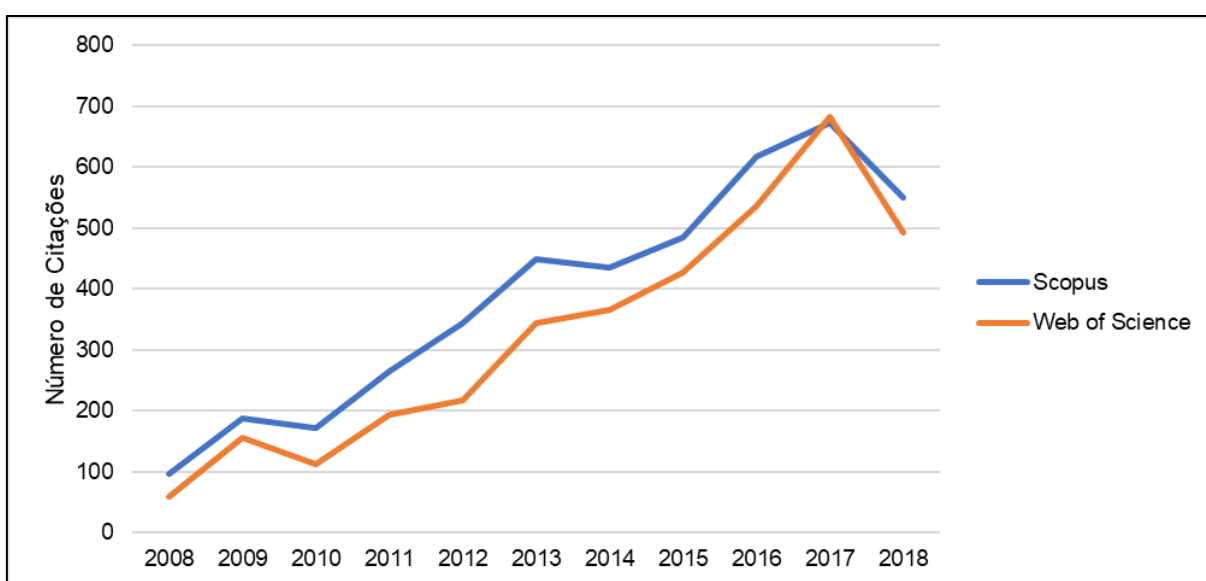


Figura 3.4 - Número de citações por ano dos documentos nas bases de dados Scopus e Web of Science

Dentre todos os documentos, os sete mais citados em cada base de dados encontram-se descritos na Tabela 3.1. Nesta tabela é possível verificar, além do título do documento, as informações de autor, ano, periódico e também o número de citações em cada base de dados.

Tabela 3.1 - Ranking dos sete documentos mais citados nas bases de dados Scopus e Web of Science

Referência (Autor/Ano)	Título da Publicação	Periódico onde foi submetido	Ranking / Número de citações	
			Scopus	WoS
Chang, T. J.; et al. (2000)	<i>Heuristics for cardinality constrained portfolio optimisation</i>	Computers & Operations Research	1° / 376	---
Doerner, K.; et al. (2004)	<i>Pareto Ant Colony Optimization: A Metaheuristic Approach to Multiobjective Portfolio Selection</i>	Annals of Operations Research	2° / 227	1° / 178
Lin, C.; Liu, Y. (2008)	<i>Genetic algorithms for portfolio selection problems with minimum transaction lots</i>	European Journal of Operational Research	3° / 133	6° / 98
Xia, Y.S.; et al. (2000)	<i>A model for portfolio selection with order of expected returns</i>	Computers & Operations Research	4° / 131	4° / 119
Li, X.; Qin, Z.; Kar, S. (2010)	<i>Mean-variance-skewness model for portfolio selection with fuzzy returns</i>	European Journal of Operational Research	5° / 127	5° / 108
Helmbold, D.P.; et al. (1998)	<i>Online Portfolio Selection using multiplicative updates</i>	Mathematical Finance	6° / 114	---
Huang, X. (2008)	<i>Mean-semivariance models for fuzzy portfolio selection</i>	Journal of Computational and Applied Mathematics	7° / 100	7° / 90
Gabrel, V.; et al. (2014)	<i>Recent advances in robust optimization: an overview</i>	European Journal of Operational Research	---	2° / 158
Steuer, R.E.; Na, P. (2003)	<i>Multiple criteria decision making combined with finance: a categorized bibliographic study</i>	European Journal of Operational Research	---	3° / 153

Como pode-se observar nesta Tabela 3.1, o documento mais citado na base Scopus é dos autores Chang, *et al.* (2000), com 376 citações, enquanto que na base Web of Science o primeiro lugar é ocupado pelos autores Doerner, *et al.* (2004), com um número de 178 citações.

Observa-se que houve pouca distinção dos documentos que entraram no *ranking* dos sete mais citados, considerando-se as bases de dados Scopus e Web Of Science, sendo que cinco dos sete documentos coincidiram em estar no *ranking* simultaneamente em ambas as bases de dados: Doerner, *et al.* (2004); Lin e Liu (2008); Xia, *et al.* (2000); Li, Qin e Kar (2010) e Huang (2008).

A proporção no número de citações destes sete trabalhos variou no decorrer dos últimos 10 anos, como mostra a Figura 3.5. Na base de dados Scopus, há uma predominância do trabalho de Chang, *et al.* (2000) desde 2008 até 2018, enquanto que na base Web of

Science o trabalho de Gabrel, *et al.* (2014) vem ganhando importância significativa desde o ano que foi publicado até o presente momento.

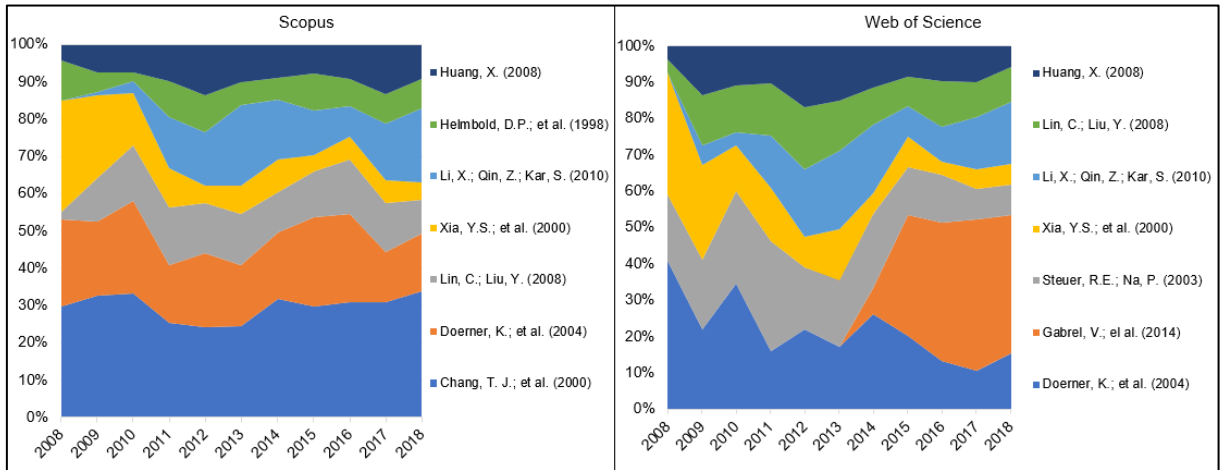


Figura 3.5 - Evolução dos sete documentos mais citados nos últimos 10 anos, nas bases de dados Scopus e Web of Science

A análise de citações, como realizada acima, é fundamental para identificar os trabalhos e autores que têm mais influência sobre o tema, uma vez que o número de citações é um indicador de grande impacto para quantificar a importância das pesquisas realizadas. Além da análise de citações, outra métrica relevante é o número de documentos por autor, como mostra a Figura 3.6, pois identifica os autores/pesquisadores mais envolvidos e que, conseqüentemente, possuem maior conhecimento e domínio do tema.

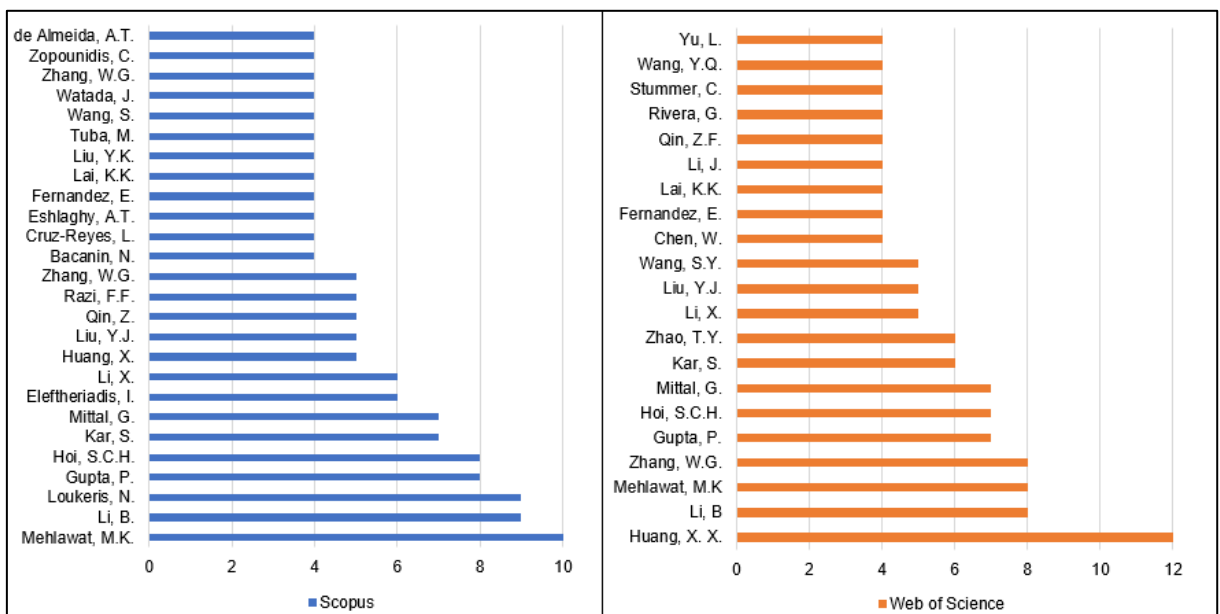


Figura 3.6 - Autores com maior número de documentos, nas bases de dados Scopus e Web of Science

Para a construção dos gráficos do número de documentos por autor (Figura 3.6), foram considerados aqueles autores com quatro ou mais artigos publicados, em cada base de dados. Os pesquisadores Huang, X. X. e Mehlawat, M. K. estão entre os autores com maior número de documentos publicados nas bases Web of Science e Scopus, tendo cada um o total de 12 e 10 artigos, respectivamente. Li, B. encontra-se em segundo lugar, em ambas as bases de dados, sendo autor/coautor de oito e nove artigos publicados. Além destes autores, outros renomados pesquisadores da área de decisão multicritério apareceram na análise, como Gupta, P; Zopounidis, C; e Almeida, A. T.

Estes pesquisadores são, em sua maioria, lotados em importantes instituições de ensino, como mostra a Figura 3.7, dentre as quais destacam-se: *Chinese Academy of Sciences*; *South China University of Technology*; *University of Delhi*; e *Islamic Azad University*.

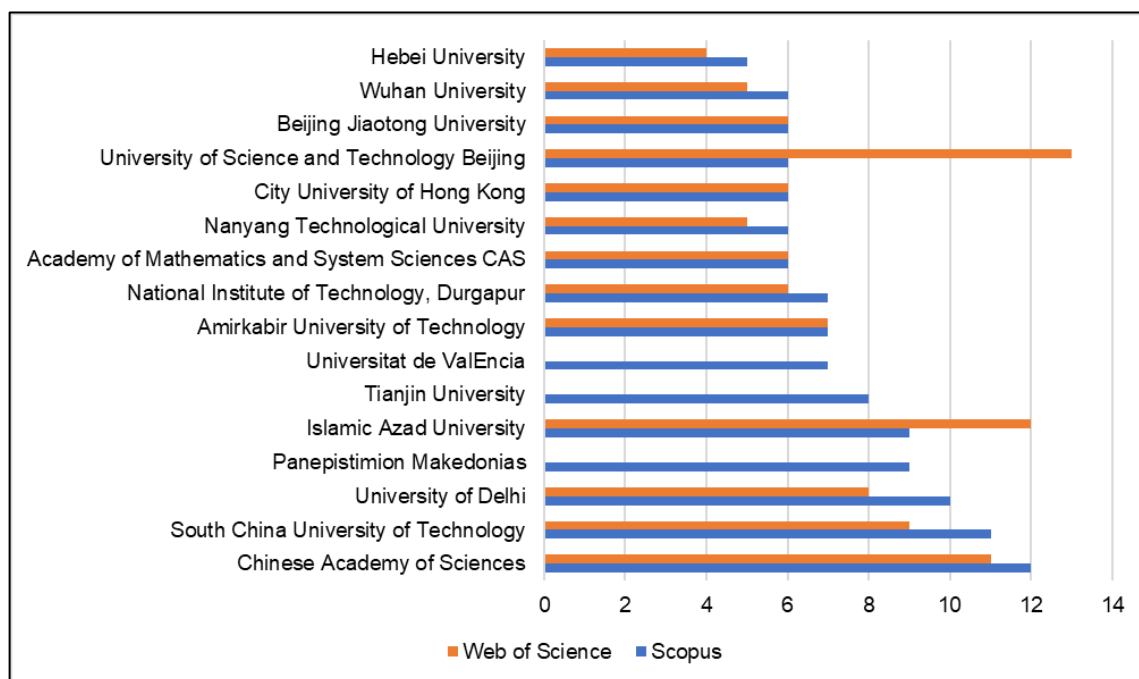


Figura 3.7 - Instituições com maior número de documentos, nas bases de dados Scopus e Web of Science

Como pode-se perceber, as instituições com pesquisadores e/ou grupos de pesquisas que se dedicam ao estudo do tema sobre seleção de portfólio estão em maior número localizadas na China, fato este que, conseqüentemente, proporciona a este país o maior número de publicações de artigos sobre o tema, como mostra a Figura 3.8. Novamente, os gráficos mostram aqueles países com quatro ou mais artigos publicados.

A China conta com um total de 115 e 108 documentos publicados, nas bases de dados Scopus e Web of Science, respectivamente. Em seguida, aparece o Irã, com aproximadamente 42 documentos, e os Estados Unidos da América, com 38 documentos publicados. Na

listagem de documentos publicados, o Brasil possui 9 e 6 documentos, nas bases Scopus e Web of Science, respectivamente.

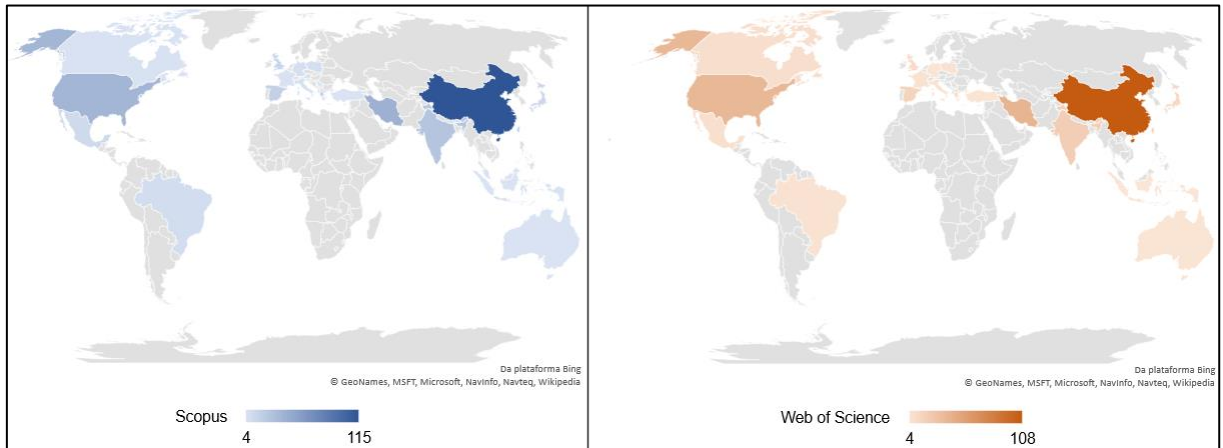


Figura 3.8 - Países com maior número de documentos, nas bases de dados Scopus e Web of Science

Na Tabela 3.1 foram elencados os periódicos dos artigos mais citados, sendo que quatro deles estão publicados no “*European Journal of Operational Research*”. Este importante periódico da área de Engenharia de Produção/PO também está entre os três periódicos com maior número de documentos publicados, como mostra a Figura 3.9. Destacam-se também os periódicos “*Expert Systems with Aplication*” e “*Lecture notes in Computer Sciences*”

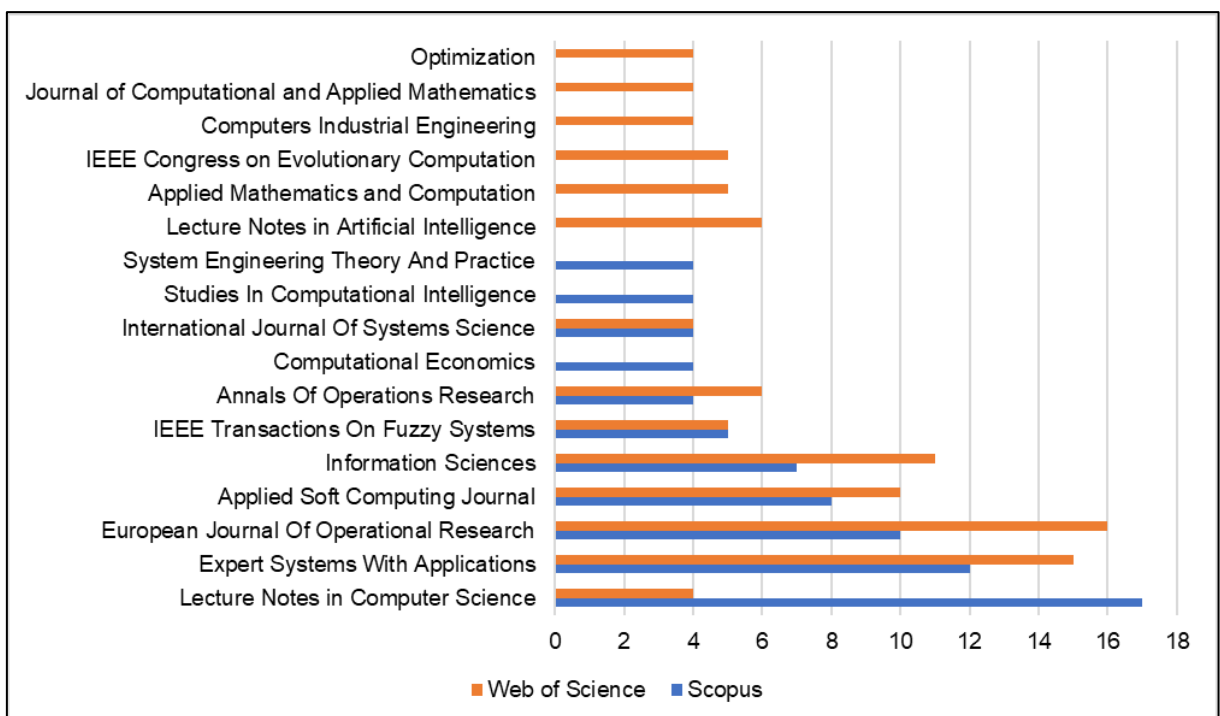


Figura 3.9 - Periódicos com maior número de documentos, nas bases de dados Scopus e Web of Science

O fato de vários dos documentos encontrados na busca serem referentes a publicações em importantes periódicos da área, com alto fator de impacto e/ou indicador Qualis/CAPES A1, demonstra a relevância e importância do tema abordado na presente pesquisa. Tais periódicos estão, geralmente, enquadrados em áreas do conhecido relativas à ciência da computação, engenharia e PO.

A Figura 3.10 mostra que a maioria dos artigos encontrados se situam nestas áreas do conhecimento, sendo que na base de dados Scopus o maior número de artigos estão nas áreas “Computer Science”, “Mathematics”, “Engineering” e “Decision Sciences”, enquanto que na base Web of Science as principais áreas onde enquadram-se os artigos são “Computer Science”, “Engineering” e “Operations Research Management Science”.

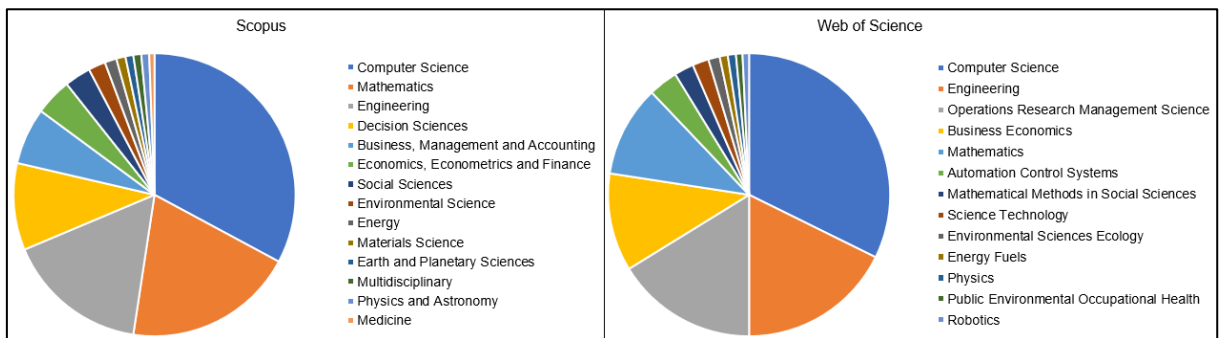


Figura 3.10 - Áreas do conhecimento das bases Scopus e Web of Science mais significativas

Fazendo uma listagem geral de todos os títulos e palavras-chave dos artigos encontrados, foi possível elaborar uma nuvem de palavras (Figura 3.11), que aponta aquelas palavras que aparecem com maior frequência, ou seja, quanto maior o tamanho da fonte, mais vezes a respectiva palavra aparece na lista.



Figura 3.11 - Nuvens de palavras com títulos e palavras-chave dos documentos encontrados

Como esperado, as palavras “*portfolio*” e “*selection*” aparecem com maior frequência entre as palavras de título e/ou palavras-chave dos artigos encontrados, como mostra a primeira nuvem de palavras da Figura 3.11. Porém, ao aproximar o *zoom* da nuvem, de forma a excluir as maiores palavras, aparecem também algumas outras palavras relevantes ao tema, tais como: *algorithm*; *genetic*; *fuzzy*; *model*; *decision*; *project*; *hybrid* e *problem*.

3.2. IDENTIFICAÇÃO DE CONCEITOS, MÉTODOS E APLICAÇÕES

Como já mencionado anteriormente, após a remoção dos artigos duplicados restaram 480 documentos, dos quais foram sistematizados e analisados quanto ao título e resumo, para a identificação dos conceitos, métodos e aplicações.

Os principais conceitos abordados pelos artigos estão relacionados com “Algoritmo Genético”, “Abordagem Multicritério” e “Teoria de conjuntos Fuzzy”, como mostra a Figura 3.12. Outros conceitos igualmente importantes, mas em menor número, também foram levantados pelos autores nas pesquisas, tais como: outras metaheurísticas (além dos Algoritmos Genético); aprendizagem de máquina; e algoritmos evolucionários multiobjetivos.

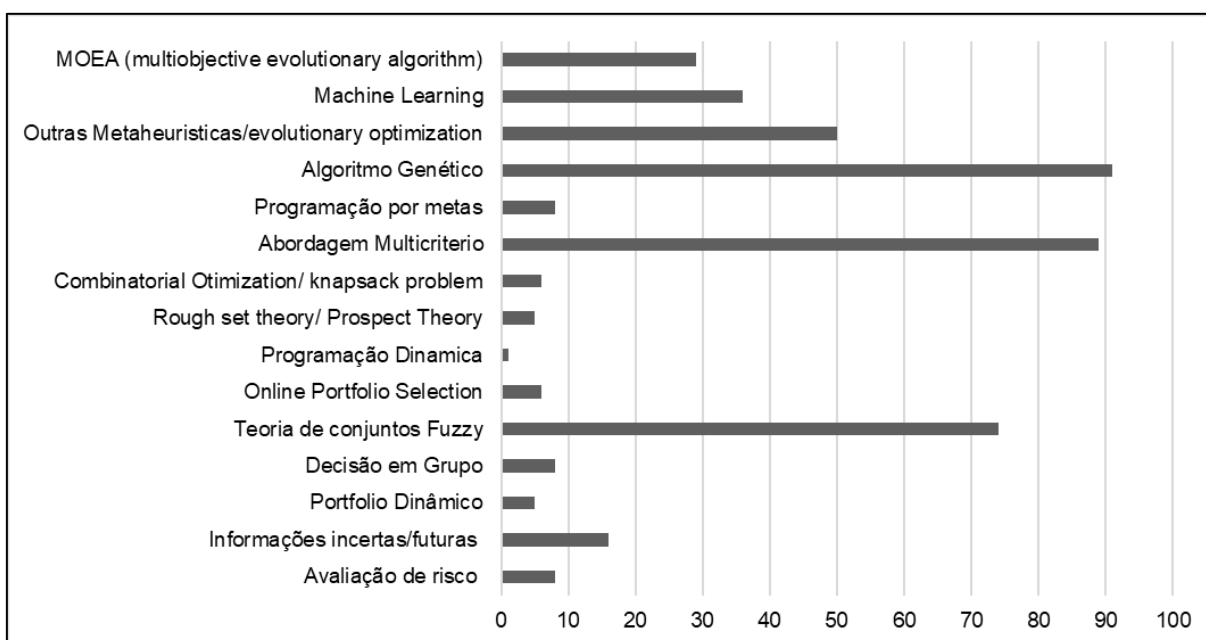


Figura 3.12 - Conceitos e abordagens elencados nos documentos encontrados

Considerando a identificação dos métodos utilizados pelos autores, estes foram classificados em três grupos principais: a) métodos com abordagem multicritério (*Multicriteria Decision Methods*, MCDM); b) métodos metaheurísticos; e c) métodos com

abordagem de aprendizagem de máquina (*Machine Learning*, ML). Também foram identificados alguns outros métodos que não se enquadram em nenhuma das três classes descritas. Na Figura 3.13 é possível verificar que as metaheurísticas compreenderam técnicas mais utilizadas pelos autores nas pesquisas sobre seleção de portfólio, seguidas dos métodos de decisão multicritério.

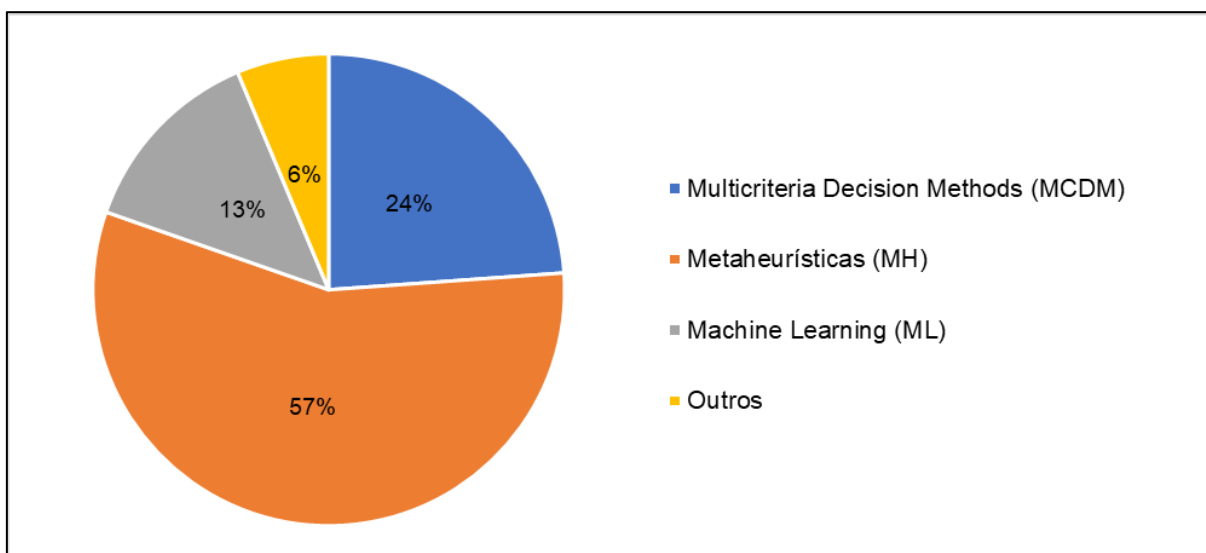


Figura 3.13 - Principais áreas dos métodos utilizados nos documentos

Entre as metaheurísticas utilizadas, mostradas na Figura 3.14, obviamente o Algoritmo Genético aparece como a técnica com maior número de documentos que o utilizam, tanto pelo fato de a busca estar direcionada especificamente a esta metaheurística, quanto pelo fato de ele ser uma técnica usualmente utilizada e difundida em várias aplicações.

Em mais de 80 documentos encontrados, os autores utilizaram-se dos algoritmos genéticos para os mais diversos fins, integrando-os a outros conceitos, tais como seleção de portfólio com abordagem *fuzzy* (Abiyev e Menekay, 2007; Barak et al., 2013; Dastkhan et al., 2011;), ou então seleção de portfólio considerando risco (Golmohammadi e Pajoutan, 2011).

Outro algoritmo metaheurístico bastante popular para a otimização multiobjetivo, que compõe a classe de algoritmos evolucionários, o NSGA II (*Non-dominated Sorting Genetic Algorithm*) também foi amplamente utilizado nas pesquisas para seleção de portfólio (Anagnostopoulos e Mamanis, 2010; Anagnostopoulos e Mamanis, 2011; Cvörnjek et al., 2014; Doerner et al., 2004; Eftekharian et al., 2017; Kar et al., 2018; Khalidji et al., 2009; Khalili-Damghani et al., 2013; Liagkouras e Metaxiotis, 2015; Martínez-Vega et al., 2018; Mashayekhi e Omrani, 2016; Nejad, 2014; Radiukyniene e Zilinskas, 2008; Razi, 2014; Razi et al., 2014; Sirisrisakulchai et al., 2015; e Wu et al., 2018)

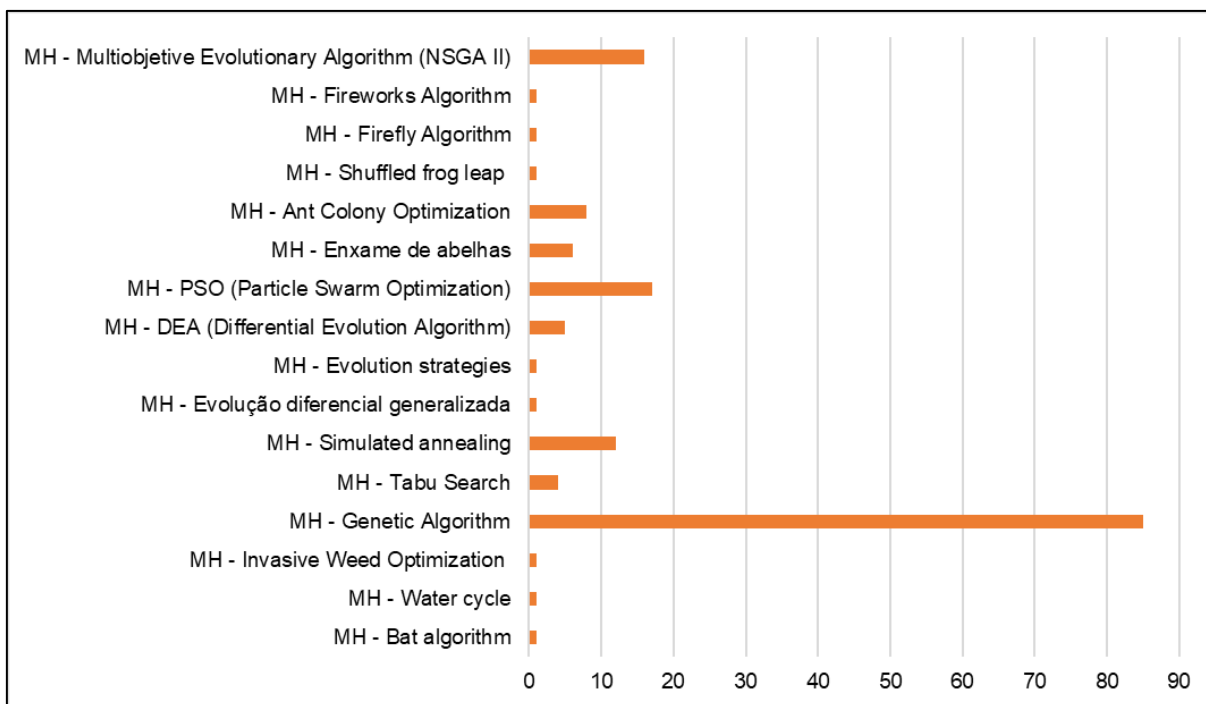


Figura 3.14 - Número de documentos que empregaram os métodos relacionados a Metaheurísticas

Os algoritmos inspirados em inteligência de enxame (também chamados de *bio-inspired algorithm*) são desenvolvidos computacionalmente observando-se a inteligência que emerge do comportamento de grupos da natureza. Alguns dos mais famosos algoritmos baseados neste tipo de inteligência também foram aplicados em problemas de seleção de portfólio, dos quais pode-se citar:

- Enxame de partícula (*Particle Swarm Optimization*, PSO; Corazza *et al.*, 2012; Deng e Chen, 2007; Deng *et al.*, 2012; Farzi *et al.*, 2013; Gilli e Schumann, 2012; Golmakani e Fazel, 2011; Kang *et al.*, 2015; Kuo e Hong, 2013; Najafi e Mushakhian, 2015; Nejad, 2014; Pouya *et al.*, 2016; Talebi *et al.*, 2010; Tofighian *et al.*, 2018; Wang *et al.*, 2011; Xu e Chen, 2009; Yaakob e Watada, 2010; Yan *et al.*, 2007);
- Colônia de formigas (Ant Colony Optimization, ACO; Cruz *et al.*, 2013; Cruz *et al.*, 2014; Doerner *et al.*, 2004; Fernandez *et al.*, 2015; Fernandez *et al.*, 2017; Khalidji *et al.*, 2009; Nejad, 2014; Tofighian e Naderi, 2015);
- Enxame de Abelhas (*Bee Colony Algorithm*, BCA; Bacanin *et al.*, 2014; Gao *et al.*, 2018; Seyedhosseini *et al.*, 2016; Sun e Wei, 2013; Tuba e Bacanin, 2014; Zuloaga e Moser, 2017)

Outras metaheurísticas utilizadas mais frequentemente pelos autores no problema de seleção de portfólio foram: *Simulated Annealing* (Chang *et al.*, 2000; Chien e Huynh, 2018;

Doerner et al., 2004; Gilli e Schumann, 2012; Gomez et al., 2006; Jahan e Akbarzadeh-Totonchi, 2010; Jahan e Akbarzadeh-T, 2012; Li, C. et al., 2018; Ruiz-Torrubiano e Suarez, 2010; Shayanfar et al., 2016; Woodside-Oriakhi et al., 2011); *Tabu Search* (Chang et al., 2000; Gilli e Schumann, 2012; Shayanfar et al., 2016; Woodside-Oriakhi et al., 2011); e *Differential Evolution Algorithm* (Andriosopoulos e Nomikos, 2014; Gilli e Schumann, 2012; Grishina et al., 2017; Liu et al., 2018; Shakhshi-Niaei et al., 2015).

Além destes, outras metaheurísticas não tão conhecidos também foram identificados, sendo considerados promissores pelos autores que as utilizaram, tais como: *Fireworks Algorithm* (Bacanin et al., 2015); *Firefly Algorithm* (Tuba e Bacanin, 2014); *Shuffled Frog-leaping Algorithm* (Razi e Shahabi, 2016); *Water Cicle Algorithm* (Moradi et al., 2017); e *Bat Algorithm* (Strumberger et al., 2016).

Já entre os métodos com abordagem de decisão multicritério (MCDM), o AHP (*Analytic Hierarchy Process*), o TOPSIS (*Technique for Order of Preference by Similarity to Ideal Solution*), o PROMETHEE (*Preference Ranking Organization Method for Enrichment of Evaluations*) e o MAUT (*Multi-Attribute Utility Theory*) foram os métodos identificados com maior frequência de aplicação nas pesquisas dos artigos encontrados na busca, como mostra a Figura 3.15.

O método AHP foi aplicado nas pesquisas de Wu et al. (2018); Patari et al. (2018); Yuen (2017); Drissi et al. (2017); Solimanpur et al. (2015); An, et al. (2015); Bahloul e Abid (2013); Mehlawat (2013); Rebiasz et al. (2013); Jiang et al. (2011); Strang (2011) e Gupta et al. (2010). Já o método TOPSIS foi utilizado por Patari et al. (2018); Yaakob et al. (2018); Drissi et al. (2017); Mishra et al. (2017); Yaakob et al. (2017); Liao e Fang (2016); Bilbao-Terol et al. (2014); Joshi e Kumar (2014); Rebiasz et al. (2013); Liu et al. (2012); e Andreica et al. (2010).

Os métodos da família PROMETHEE foram utilizados por Wu et al. (2018); Ronyastra et al. (2015); Almeida et al. (2014); López e Almeida (2014); Razi et al. (2014); Vetschera e Almeida (2012); Shakhshi-Niaei et al. (2011) e Bilsel (2007). E o MAUT foi identificado em Watson et al. (2018); Lopes e Almeida (2015); Frank et al. (2013); Sola e Mota (2012) e Yu et al. (2012).

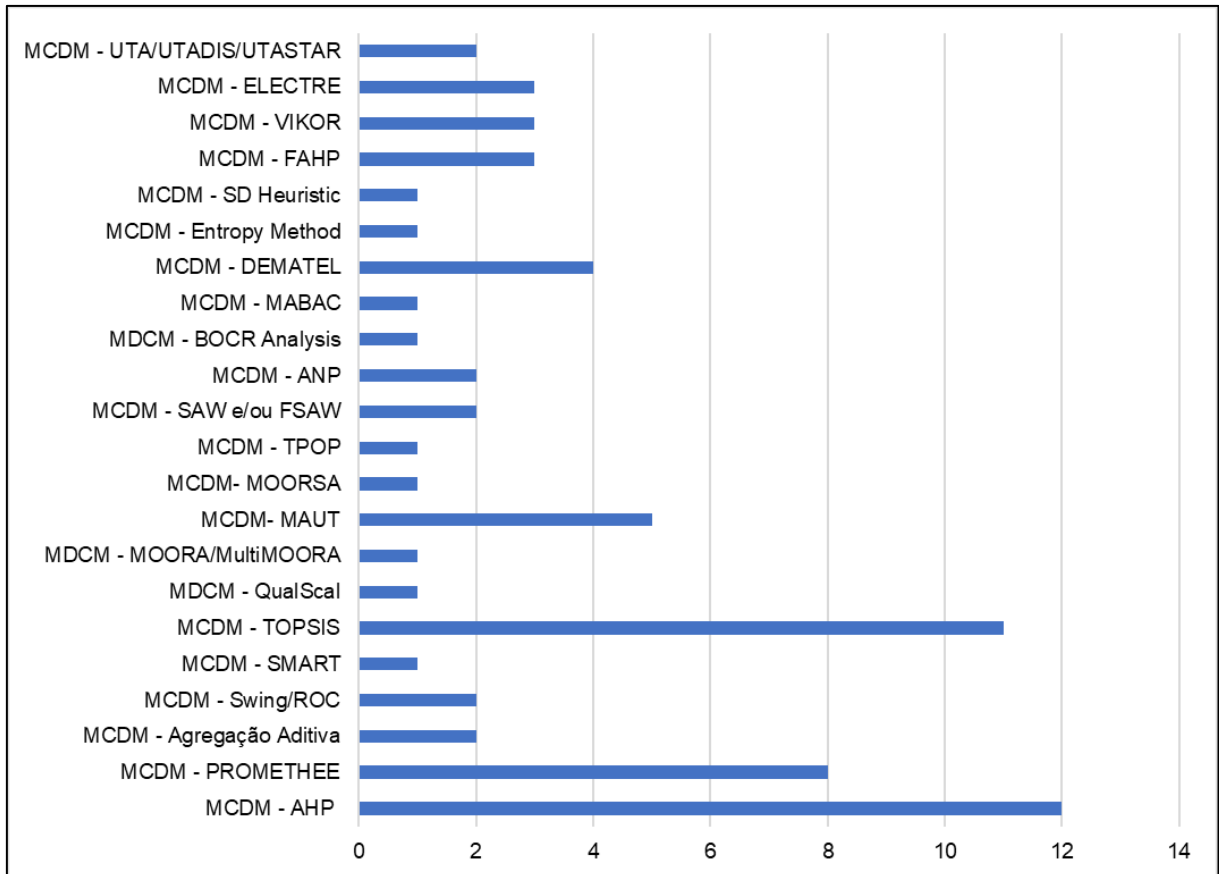


Figura 3.15 - Número de documentos que empregam os métodos relacionados a Decisão Multicritério

Além destes, outros métodos comumente conhecidos da abordagem de decisão multicritério também foram identificados nos artigos com aplicações em problemas de seleção de portfólio, tais como: DEMATEL (*Decision Making Trial and Evaluation Laboratory*; Abdollahi et al., 2015; Debnath et al., 2017; Varma e Kumar, 2012; Yang et al., 2016); ELECTRE (*Elimination et Choix Traduisant la Réalité*; Andreica et al., 2010; Peng et al., 2018; Zheng et al., 2011); VIKOR (An, et al., 2015; Drissi et al., 2017; Fouladgar et al., 2012); Fuzzy AHP (Drissi et al., 2017; Fouladgar et al., 2012; Kiris e Ustun, 2010); UTA/UTADIS (*Utilités Additives/Utilites Additives Discriminantes*; Mastorakis e Siskos, 2016; Zopounidis et al., 1999); ANP (*Analytic Network Process*; Abdollahi et al., 2015; Smith-Perera et al., 2010); SAW (*Simple Additive Weighting*) e Fuzzy SAW (Abbasianjahromi e Rajaie, 2012; Drissi et al., 2017); os procedimentos de Swing e ROC (*Rank Order Centroid*), utilizados pelos Métodos SMART/SMARTS/SMARTER (*Simple Multi-attribute Rating Technique*; Lopes e Almeida, 2013; Lopes e Almeida, 2015; Cruz-Reyes et al., 2013).

Os métodos de aprendizagem de máquina, supervisionados e não supervisionados, também foram utilizados por vários autores nas pesquisas sobre seleção de portfólio, como

mostra a Figura 3.16, dentre os quais destacam-se: a) as Redes Neurais Artificiais, sendo elas do tipo *Multi-Layer Perceptron* ou então *Radial Basis Function*; b) Técnicas não supervisionadas, como *clustering*; e c) *Support Vector Machine*.

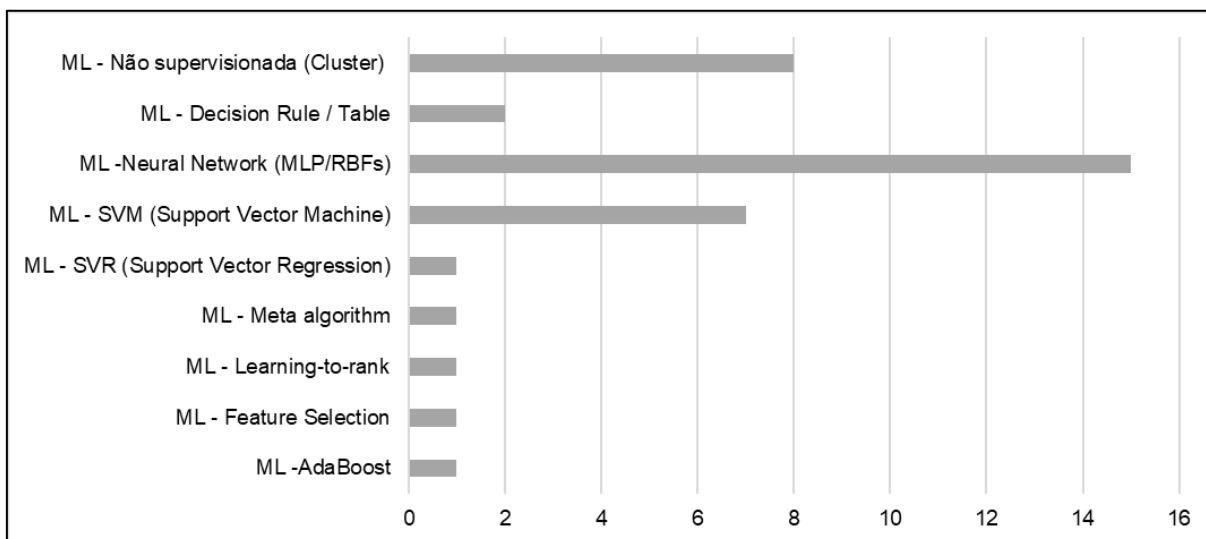


Figura 3.16 - Número de documentos que empregaram os métodos relacionados a Aprendizagem de Máquina

Os autores que aplicaram ou utilizaram cada uma das técnicas de aprendizagem de máquina, descritas na Figura 3.16, em problemas de seleção de portfólio, estão relacionados na Tabela 3.2.

Tabela 3.2 - Autores que utilizaram técnicas/métodos de aprendizagem de máquina nos estudos realizados

Técnica de Aprendizagem de Máquina	Aplicação pelos autores
<i>Clustering</i>	Ahmadi e Alizadeh Shani (2016); Boginski et al. (2014); Gupta et al. (2010); Huang e Jane (2008); Long et al. (2014); Nejad (2014); Razi et al. (2014); Razi et al. (2015)
<i>Decision Rule/Decision Table</i>	Cruz-Reyes et al. (2015); Razi e Shariat (2017); Dombi e Zsiros (2005)
<i>Artificial Neural Networks</i>	Li, C. et al. (2018); Li et al. (2009); Loukeris e Eleftheriadis (2015); Loukeris et al. (2014); Loukeris et al. (2016); Loukeris, Boutalis, Livanis, et al. (2016); Loukeris, Bekiros, et al. (2016); Prelipcean et al. (2008); Razi e Shariat (2017); Relich (2016); Roodposhti et al. (2016); Tang et al. (2004); Wang e Ieee (2006); Zhang et al. (2016).
<i>Support Vector Machine</i>	Gupta et al. (2012); Gupta et al. (2014); Huang et al. (2011); Loukeris et al. (2013); Loukeris et al. (2016); Loukeris, Eleftheriadis, et al. (2016); Paiva et al. (2018)
<i>Support Vector Regression</i>	Takeda et al. (2010)
<i>Meta Algorithm</i>	Das e Banerjee (2011)
<i>Learning to rank</i>	Song et al. (2017)
<i>Feature Selection</i>	Razi e Shariat (2017)
<i>AdaBoost</i>	Huang et al. (2011)

E, finalizando a análise de métodos e técnicas aplicadas em problemas de seleção de portfólio, alguns autores utilizaram técnicas relevantes, mas que não se enquadram em nenhuma das classificações feitas anteriormente, como é o caso da Análise Envoltória dos Dados (*Data Envelopment Analysis*, DEA), um método estatístico não paramétrico, muito utilizado em PO e economia para a estimativa de fronteiras de produção, que também foi utilizada por Abdollahi et al. (2015); Khalili-Damghani et al. (2013); Kuo e Hong (2013); Mashayekhi e Omrani (2016); Patari et al. (2018); Pouya et al. (2016) e Tavana et al. (2013) para análise em problemas de seleção de portfólio.

A Figura 3.17 apresenta esta e outras técnicas/métodos utilizados pelos autores, tais como a *Gray Relational Analysis* (Razi, 2014; Razi e Shariat, 2017; Razi et al., 2015; Razi e Shahabi, 2016); *Failure Mode and Effect Analysis* (Foroozesh et al., 2018); Análise SWOT (*Strengths, Weaknesses, Opportunities, Threats*; Dehouche, 2015 e Frank et al., 2013); e a simulação de Monte Carlo, aplicada a programação dinâmica (Detemple, 2014; Ronyastra et al., 2015; Jahan e Akbarzadeh-T, 2012; Shakhshi-Niaei et al., 2011).

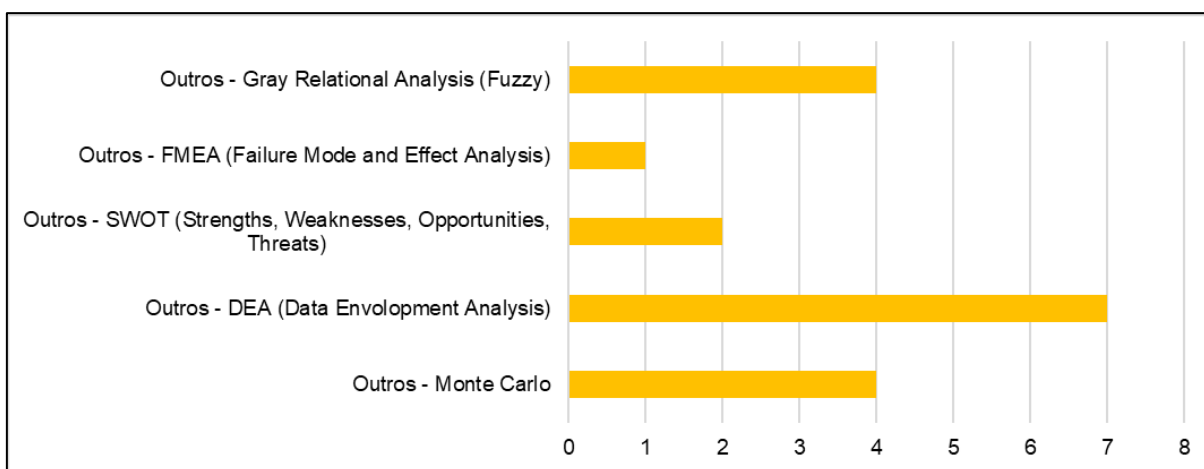


Figura 3.17 - Número de documentos que empregaram outros métodos

Além da identificação dos métodos/técnicas utilizados pelos autores para a resolução de problemas relacionados à seleção de portfólio, com a análise de título e resumo de cada um dos artigos também foi possível identificar a aplicação feita por eles, podendo esta ser um estudo de caso em empresas reais, uma aplicação com informação de banco de dados, ou então apenas a exemplificação com um problema numérico hipotético.

A Figura 3.18 demonstra a quantidade de artigos em cada uma das aplicações, sendo predominante os estudos que utilizaram informações dos índices de bolsa de valores ou mercado de ações, tais como exemplo os ativos da: *Bombay Stock Exchange* (Bhattacharyya, Kar, et al., 2011; Bhattacharyya et al., 2014; Dutta et al., 2018); *Istanbul Stock Exchange*

(Bilsel, 2007); *Spanish Stock Exchange* (Cadenas et al., 2012; Vercher e Bermudez, 2015); *Taiwan Stock Exchange* (Chen et al., 2013; Huang e Jane, 2008; Kang et al., 2015); *New York Stock Exchange* (Dastkhan et al., 2011); *Shanghai Stock Exchange* (Deng e Li, 2014; Magoc e Modave, 2011; Yu et al., 2009); *Tehran Stock Exchange* (Kimiagari et al., 2010; Najafi e Mushakhian, 2015; Nejad, 2014; Seyedhosseini et al., 2016; Soureh e Amanollahi, 2017; Talebi et al., 2010); *Chinese Stock Exchange* (Liu et al., 2012; Liu e Zhang, 2015); *Thailand Stock Exchange* (Long et al., 2014); *Athens Stock Exchange* (Zopounidis et al., 1999); e *São Paulo Stock Exchange* (Andriosopoulos e Nomikos, 2014; Paiva et al., 2018).

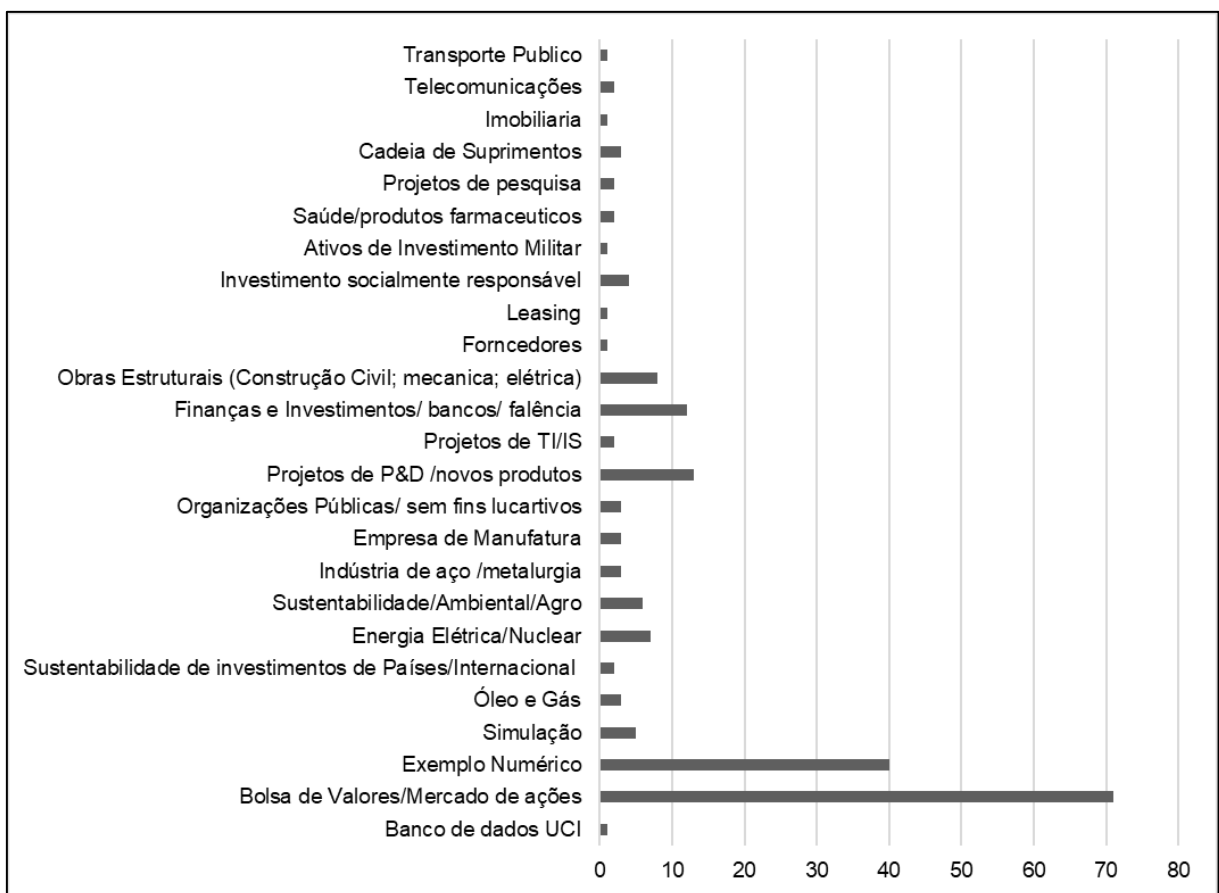


Figura 3.18 - Campos de aplicações práticas nos quais foram realizados estudos

A exemplificação com um problema numérico hipotético também foi bastante usada pelos autores, muitas vezes para simplificar a utilização e/ou eficiência de um novo método/técnica proposta (Ali et al., 2017; Andreica et al., 2010; Barak et al., 2013; Barati et al., 2016; Cadenas et al., 2015; Calvo et al., 2014; Doerner et al., 2004; Emmerich et al., 2016; Feng et al., 2017; Fernandez et al., 2015; Gao et al., 2018; Guo et al., 2016; Hassanlou, 2017; Huang et al., 2016; Huang e Zhao, 2014; Li e Xu, 2013; Martínez-Vega et al., 2018; Rabbani et al., 2010; Razi, 2014; Takeda et al., 2010; Tang et al., 2004; Tchangani e Pérès,

2017; Tofighian e Naderi, 2015; Vernic, 2016; Wang et al., 2011; Xia et al., 2000; Xia et al., 2001; Yan et al., 2007; Zhang, W. G. et al., 2011; Zheng et al., 2011).

Além disso, outras aplicações em estudo de casos reais bastante interessantes foram realizadas pelos autores em diferentes áreas, tais como em empresas de óleo e gás (Lopes e Almeida, 2013; Lopes e Almeida, 2015; Sirisrisakulchai et al., 2015), energia elétrica ou energia nuclear (Almeida e Duarte, 2011; Hocine et al., 2018; López e Almeida, 2014; Smith-Perera et al., 2010; Strang, 2011; Watson et al., 2018; Wu et al., 2018), indústrias de aço e metalurgia (Rebiasz et al., 2013; Rebiasz et al., 2014; Vilkkumaa et al., 2018), entre outras aplicações. A relação dos autores e de todas as aplicações feitas nos diferentes ramos de atividades ou empresas, estão apresentadas na Tabela 3.3.

Tabela 3.3 - Autores que realizaram estudos aplicados em diferentes empresas ou ramo de atividades

Ramo de atividades ou empresas	Aplicação pelos autores
Óleo e Gás	Lopes e Almeida (2013); Lopes e Almeida (2015); Sirisrisakulchai et al. (2015)
Sustentabilidade de investimentos	Bahloul e Abid (2013); Bilbao-Terol et al. (2014)
Energia Elétrica ou Nuclear	Almeida e Duarte (2011); Hocine et al. (2018); López e Almeida (2014); Smith-Perera et al. (2010); Strang (2011); Watson et al. (2018); Wu et al. (2018)
Sustentabilidade; ambiental ou Agronegócio	An, et al. (2015); Cadenas et al. (2015); Calvo et al. (2014); Debnath et al. (2017); Jiang et al. (2011); Xidonas et al. (2016)
Indústria de aço ou metalurgia	Rebiasz et al. (2013); Rebiasz et al. (2014); Vilkkumaa et al. (2018)
Empresa de Manufatura	Cho e Moon (2006); Liao e Fang (2016); Sola e Mota (2012)
Organizações Públicas/ sem fins lucrativos	Almeida et al. (2014); Cruz-Reyes et al. (2015); Dehouche (2015)
Projetos de P&D /novos produtos	Bhattacharyya, Kumar, et al. (2011); Chien e Huynh (2018); Almeida e Duarte (2011); Doerner et al. (2004); Guo et al. (2018); Huang e Zhao (2014); Lin (2007); Relich (2016); Relich e Pawlewski (2017); Stewart (1991); Stewart (2016); Stummer e Kiesling (2009); Van Dyk e Smith (1990)
Projetos de TI/SI	Almeida et al. (2014); Joshi e Kumar (2014); Lin (2007)
Finanças e Investimentos/ bancos/ falência	Arasteh et al. (2014); Engau (2009); Gilli e Schumann (2012); Joshi e Kumar (2014); Loukeris et al. (2013); Loukeris et al. (2014); Loukeris e Eleftheriadis (2015); Loukeris, Boutalis, et al. (2016); Loukeris, Eleftheriadis, et al. (2016); Metawa et al. (2017); Solimanpur et al. (2015); Zhang et al. (2011)
Obras Estruturais (Construção Civil; mecânica; elétrica)	Abbasianjahromi e Rajaie (2012); Bhattacharyya, Kumar, et al. (2011); Dobrovolskiene e Tamosiuniene (2016; 2016b); Dorfeshan et al. (2018); Razi et al. (2014); Razi e Shariat (2017); Shayanfar et al. (2016)
Fornecedores	Abdollahi et al. (2015)
Leasing	Andreica et al. (2010)
Investimento socialmente responsável	Bilbao-Terol et al. (2012); Cadenas et al. (2015); Calvo et al. (2014); Cruz-Reyes et al. (2013)
Investimento Militar	Chang et al. (2010)
Saúde/produtos farmacêuticos	Convertino e Convertino (2014); Mastorakis e Siskos (2016)
Projetos de pesquisa	Fernandez et al. (2013); Fernandez et al. (2015)
Cadeia de Suprimentos	Foroozesh et al. (2018); Liao e Fang (2016); Zhang et al. (2016)
Imobiliária	Ronyastra et al. (2015)
Telecomunicações	Shakhsi-Niaei et al. (2015); Shakhsi-Niaei et al. (2011)
Transporte publico	Yang et al. (2016)

3.3. DESCRIÇÃO DOS TRABALHOS CORRELATOS

Pela leitura e análise do título e resumo dos 480 artigos, observou-se a divisão dos problemas de seleção de portfólio em duas categorias principais, que também é considerada por Abbasianjahromi e Rajaie (2012): no primeiro grupo, encontram-se os problemas exclusivamente de riscos financeiros, baseados na variância dos retornos dos projetos como medida de risco para a escolha ótima do portfólio, sendo o modelo de média-variância de Markowitz, e suas variações, o principal foco na resolução deste tipo de problema; e, no segundo grupo, encontram-se os problemas que são avaliados e classificados não somente com base nos seus riscos financeiros, mas também com base em outros critérios/atributos, sendo os métodos de decisão multicritério importantes ferramentas para resolução deste tipo de problema.

As aplicações práticas do primeiro grupo de problemas de seleção de portfólio são, basicamente, aquelas nas quais os autores utilizaram dados de bolsa de valores. Já no segundo grupo, encontram-se os problemas de seleção de portfólio cujas aplicações vão além da área de mercado de ações, e englobam a resolução de problemas mais amplos, considerando vários critérios, em empresas de diversos ramos.

Para a presente pesquisa, o interesse primordial está nos artigos que se enquadram no segundo grupo de problemas de seleção de portfólio, ou seja, aqueles com abordagem voltada à decisão multicritério, cujas as aplicações consideram vários critérios, não necessariamente que envolvam riscos financeiros. Considerando isso, foram selecionados 80 dos 480 artigos, que tratavam do problema de seleção de portfólio com uma visão voltada à decisão multicritério, para análise de conteúdo e, posteriormente, para a descrição dos trabalhos correlatos.

Destes 80 artigos considerados para a análise nesta etapa da revisão de literatura, 43% aplicaram um ou mais métodos já existentes; 24% propuseram novas abordagens, metodologias ou sistemas; outros 24% fizeram a integração de dois ou mais métodos como proposta de um novo *framework* de análise; e 9% são artigos de revisão de literatura, *surveys* ou discussões, como mostra a Figura 3.19.

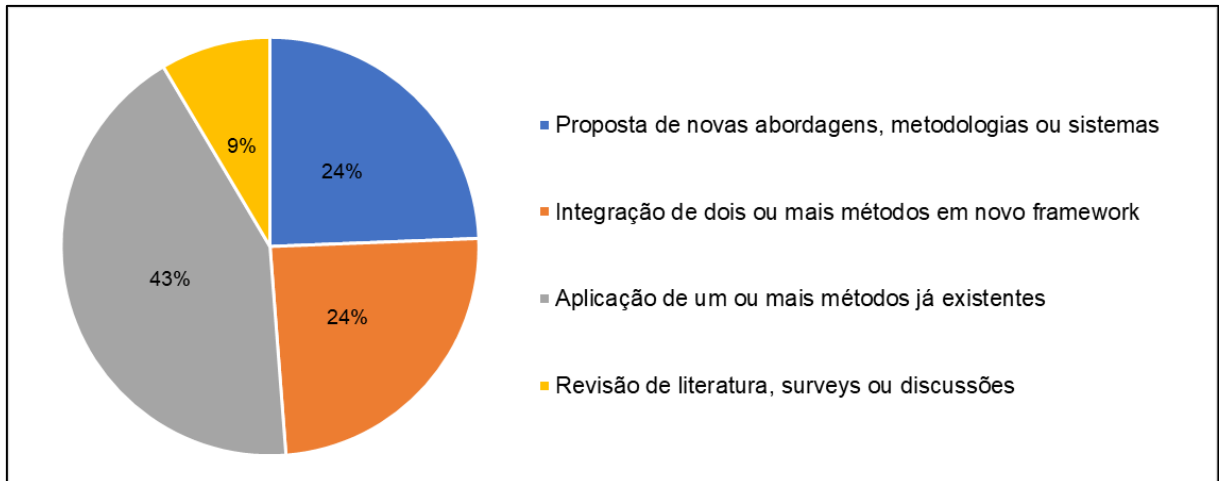


Figura 3.19 - Proporção dos tipos de pesquisa nos quais se enquadram os trabalhos correlatos

Os artigos de revisão de literatura ou *surveys* (Belaid et al., 2013; Colapinto et al., 2017; Doumpou e Zopounidis, 2014; Morton et al., 2016; Morton, 2015; Steuer e Na, 2003; Zopounidis e Doumpou, 2013) foram analisados na íntegra, pois são de suma importância, uma vez que trazem o estado da arte no desenvolvimento das pesquisas na área de seleção de portfólio. Com a leitura destes e de outros artigos mais relevantes, constatou-se que nenhum trabalho apresentou o mesmo objetivo desta pesquisa, ou seja, a proposta de um DSS englobando em um *framework* métodos e técnicas de aprendizagem de máquina, decisão multicritério e computação evolucionária.

Isso novamente enfatiza a ineditismo da pesquisa proposta, muito embora algumas semelhanças parciais tenham sido identificadas como, por exemplo, com relação ao: a) uso dos mesmos métodos: Lopes e Almeida (2013) e Lopes e Almeida (2015) utilizaram o procedimento de *swing* para determinação dos pesos dos critérios, assim como realizado nesta pesquisa para a elicitação dos pesos junto ao decisor; b) aplicação dos métodos na mesma empresa do estudo de caso: Almeida e Duarte (2011), e López e Almeida (2014) também aplicaram métodos para seleção de portfólio na empresa de energia elétrica considerada como um dos estudos de caso.

E, a seguir, estão descritas algumas das pesquisas mais relevantes encontradas na revisão de literatura. Ali *et al.* (2017) consideraram a lógica *fuzzy* para a implementação de um sistema que considera a entrada de dois critérios (taxa de retorno através da introdução de um projeto inovador; e efeitos econômicos esperados usando tecnologias mais produtivas) e, como saída, a rentabilidade do projeto. Aplicaram o modelo em um exemplo numérico simulado, cujo resultado de uma alternativa, que tem 5 de taxa de retorno e 5 de efeito

econômico, será de 5,18 de rentabilidade, e uma outra alternativa, com 100 de taxa de retorno e 90 de efeito econômico, terá 92,8 de rentabilidade.

Ahmadi e Alizadeh Shani (2016) aplicaram métodos de decisão multicritério para definição de pesos e *ranking*, juntamente com a análise de *clustering*, em um estudo de caso de companhias farmacêuticas da bolsa de valores. Foram considerados para a análise cinco critérios e sete companhias e, como resultado, indicou a seleção de três companhias, que compõe o primeiro nível do *ranking*.

A seleção de portfólio com uma abordagem sustentável foi foco do estudo de An *et al.* (2015), que integraram os métodos AHP e VIKOR para a seleção do portfólio de ações mais eficazes, de modo a resolver os graves problemas causados pela reciclagem informal de lixo eletrônico da China. Um estudo de caso ilustrativo foi aplicado na cidade de Guiyu, e evidenciou que as melhores ações incluem: o apoio aos revendedores informais para transição formal; o investimento em infraestrutura para reciclagem; o treinamento e educação da população; e a restrição de importação de resíduos de equipamentos elétricos e eletrônicos.

Abdollahi *et al.* (2015) apresentaram um *framework* para a seleção de fornecedores considerando os critérios “ágeis” (relacionado a organização) e “enxutos” (relacionado ao produto), integrando os métodos ANP, DEMATEL e DEA. Aplicaram o *framework* em um exemplo numérico, contendo três critérios “enxutos” (custo; qualidade; e entrega) e quatro critérios “ágeis” (humano; tecnologia; gerencial; e cultural), para classificação e seleção de fornecedores em uma matriz “*Low/High*” dos critérios considerados.

Bilbao-Terol *et al.* (2014) utilizam o método TOPSIS para avaliar a sustentabilidade de fundos de investimentos internacionais auxiliando, assim, investidores a tomar decisões baseado na unificação de indicadores de sustentabilidade do fundo. Foram usados indicadores como: economia líquida ajustada; pegada ecológica; índice de desempenho ambiental; e o índice de desenvolvimento humano. Também, foram considerados 105 países para a análise geral e, para uma análise mais profunda, três países foram escolhidos: França, Holanda e Itália.

O impacto das barreiras de investimentos na seleção internacional de portfólios foi foco do estudo de Bahloul e Abid (2013), que desenvolveram uma abordagem integrada de tomada de decisão multicritério, combinando o método AHP com o modelo de Programação por Metas. O AHP foi usado para determinar os portfólios internacionais adequados em relação a sete barreiras ao investimento internacional (custo de informação; comportamento do investidor; distância geográfica; custos de transação; risco de expropriação; tamanho do mercado financeiro; e restrições de fluxo de capital). E a programação por metas,

incorporando os pesos encontrados pelo AHP, foi formulada para determinar os portfólios ótimos.

Abbasianjahromi e Rajaie (2012) apresentam um modelo para a seleção de projetos, incluindo os conceitos de risco em portfólio, conjuntos *fuzzy* e decisão multicritério. Este modelo identifica o nível de risco suportável da empresa (RELC, *risk endurable level of company*), sendo uma ferramenta para a empresa determinar a taxa de risco que ela pode suportar. Também apresentam um exemplo da aplicação do modelo em um estudo de caso real, com uma empresa do ramo da construção de estradas e barragens.

No ramo de projetos de Pesquisa e Desenvolvimento (P&D), a decisão do melhor portfólio envolve eventos e oportunidades futuras, e grande parte das informações necessárias é incerta e pouco confiável. Porém, as empresas investem em projetos de P&D para que os recursos sejam focados em resultados bem-sucedidos. A seleção dos projetos mais apropriados é complicada por muitos fatores, como incerteza, interdependência entre projetos, risco e longo prazo de entrega, que são difíceis de mensurar. Por conta disso, Bhattacharyya, Kumar *et al.* (2011) apresentam uma abordagem *fuzzy* de programação multiobjetivo para facilitar a tomada de decisão na seleção de projetos de P&D, onde a solução é obtida por algoritmo genético, e algoritmo genético multiobjetivo.

Além da dificuldade na seleção de portfólio por conta de informações incertas, muitas vezes os dados são fornecidos na forma de intervalos e, para englobar este tipo de dados na análise, Andreica *et al.* (2010) estenderam os métodos TOPSIS e ELECTRE III para os casos de seleção de portfólios com informações intervalares. Para exemplificar o procedimento do algoritmo proposto, foi utilizado um problema de seleção de portfólio para uma empresa de *leasing*, além de serem realizadas diversas simulações numéricas.

4. PROPOSTA DE UM MODELO PARA ANÁLISE DE DECISÃO MULTICRITÉRIO EM PROBLEMAS DE SELEÇÃO DE PORTFÓLIO

Neste capítulo é apresentado um modelo para a análise de decisão multicritério, que também pode ser denominado como um sistema de apoio à decisão (*Decision Support System*, DSS), específico para problemas de seleção de portfólio, baseando-se em métodos e/ou técnicas de aprendizagem de máquina e de computação evolucionária, para facilitar e otimizar a definição das constantes de escala (pesos) dos critérios, em busca de melhores soluções para o problema.

Primeiramente são definidos os conjuntos de dados, bem como os vetores e matrizes de dados, necessários para a aplicação dos métodos e/ou técnicas do modelo proposto. A forma de utilização desses métodos e/ou técnicas foi estruturada em um processo sequencial, apresentado em um *framework* na seção 4.2, cujas fases são detalhadas nas seções que seguem.

4.1. DEFINIÇÃO DOS CONJUNTOS DE DADOS, VETORES E MATRIZES

Dado um problema de decisão multicritério para seleção de portfólio, com n alternativas (projetos ou instâncias) e m critérios (atributos), a relação da definição de todos os conjuntos de dados, matrizes e vetores utilizados no modelo proposto é apresentada como:

- \mathbb{A} : conjunto de alternativas (projetos ou instâncias):

$$\mathbb{A} = \{a_i | 1 \leq i \leq n\}$$

- \mathbb{C} : conjunto de critérios (atributos):

$$\mathbb{C} = \{c_j | 1 \leq j \leq m\}$$

- K : vetor de constantes de escala (pesos) para cada critério:

$$K = [k_1, \dots, k_j, \dots, k_m]$$

- \mathbb{D} : conjunto de desempenho de cada alternativa em cada critério:

$$\mathbb{D} = \{d_{ij} | 1 \leq i \leq n \text{ e } 1 \leq j \leq m\}$$

- \mathbb{B} : conjunto de custo para o desenvolvimento de cada alternativa:

$$\mathbb{B} = \{b_i | 1 \leq i \leq n\}$$

- B : orçamento total disponível (*budget*):

$$B \in R_+^*$$

- \mathbb{X} : conjunto dos projetos selecionados (ou não) para o portfólio:

$$\mathbb{X} = \{x_i \in \{0,1\} \forall i = 1 \dots n \text{ tal que } 0: \text{"não selecionado"}; 1: \text{"selecionado"}\}$$

- \mathbb{Y} : conjunto da classificação preditiva das alternativas (projetos ou instâncias):

$$\mathbb{Y} = \{y_i \in \{0,1\} \forall i = 1 \dots n | 0: \text{"insucesso"}; 1: \text{"sucesso"}\}$$

- \mathbb{A}^h : conjunto de alternativas (projetos ou instâncias) passadas (registros históricos):

$$\mathbb{A}^h = \{a_i^h | 1 \leq i \leq n\}$$

- \mathbb{D}^h : conjunto de desempenho de cada instância a_i^h em cada atributo (registros históricos):

$$\mathbb{D}^h = \{d_{ij}^h | 1 \leq i \leq n \text{ e } 1 \leq j \leq m\}$$

- \mathbb{Y}^h : conjunto da classificação das instâncias a_i^h (registros históricos):

$$\mathbb{Y}^h = \{y_i^h \in \{0,1\} \forall i = 1 \dots n \text{ tal que } 0: \text{"insucesso"}; 1: \text{"sucesso"}\}$$

Os valores das constantes de escala (pesos), definidos pelo vetor K , podem ser diferenciados conforme o modo pelo qual foram obtidos, utilizando-se das siglas sobrescritas FS, ED, AG-FS, AG-ED e AG, para simbolizar que o vetor é resultado dos procedimentos de:

- *Feature Selection*, K^{FS} ;
- Elicitação com o Decisor, K^{ED} ;
- Otimização com Algoritmo Genético (respeitando a ordem de importância/preferência dos critérios, estabelecida pelo procedimento de *Feature Selection*), K^{AG-FS} ;
- Otimização com Algoritmo Genético (respeitando a ordem de importância/preferência dos critérios, estabelecida pelo procedimento de Elicitação com Decisor), K^{AG-ED} ;

- Otimização com Algoritmo Genético (sem considerar nenhuma ordem de importância/preferência dos critérios previamente estabelecida), K^{AG} .

Os elementos dos conjuntos \mathbb{A} , \mathbb{C} , \mathbb{D} e \mathbb{B} , anteriormente descritos, juntamente com o orçamento total disponível (*budget*), B , compõem a Matriz de Avaliação inicial, MA (Quadro 4.1), e simbolizam as informações dos projetos a serem avaliados para compor (ou não) o portfólio.

	c_1	c_2	...	c_m	b :
a_1	d_{11}	d_{12}	...	d_{1m}	b_1
a_2	d_{21}	d_{22}	...	d_{2m}	b_2
...
a_n	d_{n1}	d_{n2}	...	d_{nm}	b_n
					B

Quadro 4.1 - Esquema da Matriz de Avaliação (MA).

Similarmente, caso existam registros históricos de decisões semelhantes passadas h , com os elementos dos conjuntos \mathbb{A}^h , \mathbb{C} , \mathbb{D}^h e \mathbb{B}^h , juntamente com a possível classificação dos projetos passados em “sucesso” ou “insucesso” (conjunto \mathbb{Y}^h), têm-se a Matriz de Avaliação de dados históricos, MA^h (Quadro 4.2).

	c_1	c_2	...	c_m	y^h :
a_1^h	d_{11}^h	d_{12}^h	...	d_{1m}^h	y_1^h
a_2^h	d_{21}^h	d_{22}^h	...	d_{2m}^h	y_2^h
...
a_n^h	d_{n1}^h	d_{n2}^h	...	d_{nm}^h	y_n^h

Quadro 4.2 - Esquema da Matriz de Avaliação de Dados Históricos (MA^h)

Além de representar os projetos já realizados, esta matriz se diferencia da MA por conter a informação de rótulo (*label*) dos projetos, ou seja, a qual classe (“sucesso” ou “insucesso”) cada projeto realizado pertence, e também por não necessitar das informações do custo de cada projeto e do orçamento disponível (*budget*, B).

Observe-se que o conjunto de critérios (atributos) não se altera de uma matriz para a outra, ou seja, os mesmos critérios (atributos) que são considerados para a análise presente devem ser os mesmos pelos quais os projetos foram avaliados anteriormente.

4.2. FRAMEWORK DO MODELO PROPOSTO

Como o modelo proposto envolve vários métodos e/ou técnicas, o processo de análise foi estruturado de forma a agrupar em procedimentos distintos os métodos e/ou técnicas que tenham a mesma abordagem e/ou objetivo. Por sua vez, estes procedimentos foram organizados em quatro fases principais, sendo elas: a) Fase Preliminar; b) Fase de Predição; c) Fase de Elicitação com o Decisor; e d) Fase de Seleção e Otimização.

A Tabela 4.1 apresenta um resumo dos objetivos de cada procedimento em cada fase. Estes procedimentos são detalhados nas seções 4.3 a 4.6, juntamente com os métodos e/ou técnicas possíveis de serem utilizados.

Tabela 4.1 - Resumo do objetivo de cada procedimento em cada fase do modelo proposto

Fase	Procedimento	Objetivo (s)
Preliminar	Dados Iniciais	Caracterizar o decisor; definir os critérios (atributos); estabelecer as alternativas (instâncias/projetos).
	Modelagem de Preferências	Definir a estrutura de preferência; determinar o tipo de racionalidade.
	Avaliação Intracritério	Definir a função de valor $v_j(a_i)$.
Predição	Aprendizagem Supervisionada	Reconhecimento de padrões para previsão e classificação das instâncias.
	Seleção dos Atributos	Estabelecer um <i>ranking</i> e valores para os atributos, com base nos registros históricos.
Elicitação com Decisor	Avaliação Intercriério	Estabelecer uma ordem e função valor $v(a_i)$, com base na preferência do decisor.
Seleção e Otimização	Otimização Combinatória	Selecionar os projetos que maximizem o valor do portfólio.
	Computação Evolucionária	Otimizar o vetor de constantes de escala (pesos) dos critérios (atributos).

Sendo assim, o *framework* geral do modelo proposto (Figura 4.1) é composto, além das fases descritas na Tabela 4.1, pelas informações de entrada (matrizes MA e MA^h), por duas questões (“Existem registros históricos?” e “O decisor está disponível?”), que direcionam o fluxo a ser seguido, e é finalizado com a apresentação dos resultados, que consistem nas informações dos projetos selecionados para o portfólio (x_i) e o valor do portfólio (V).

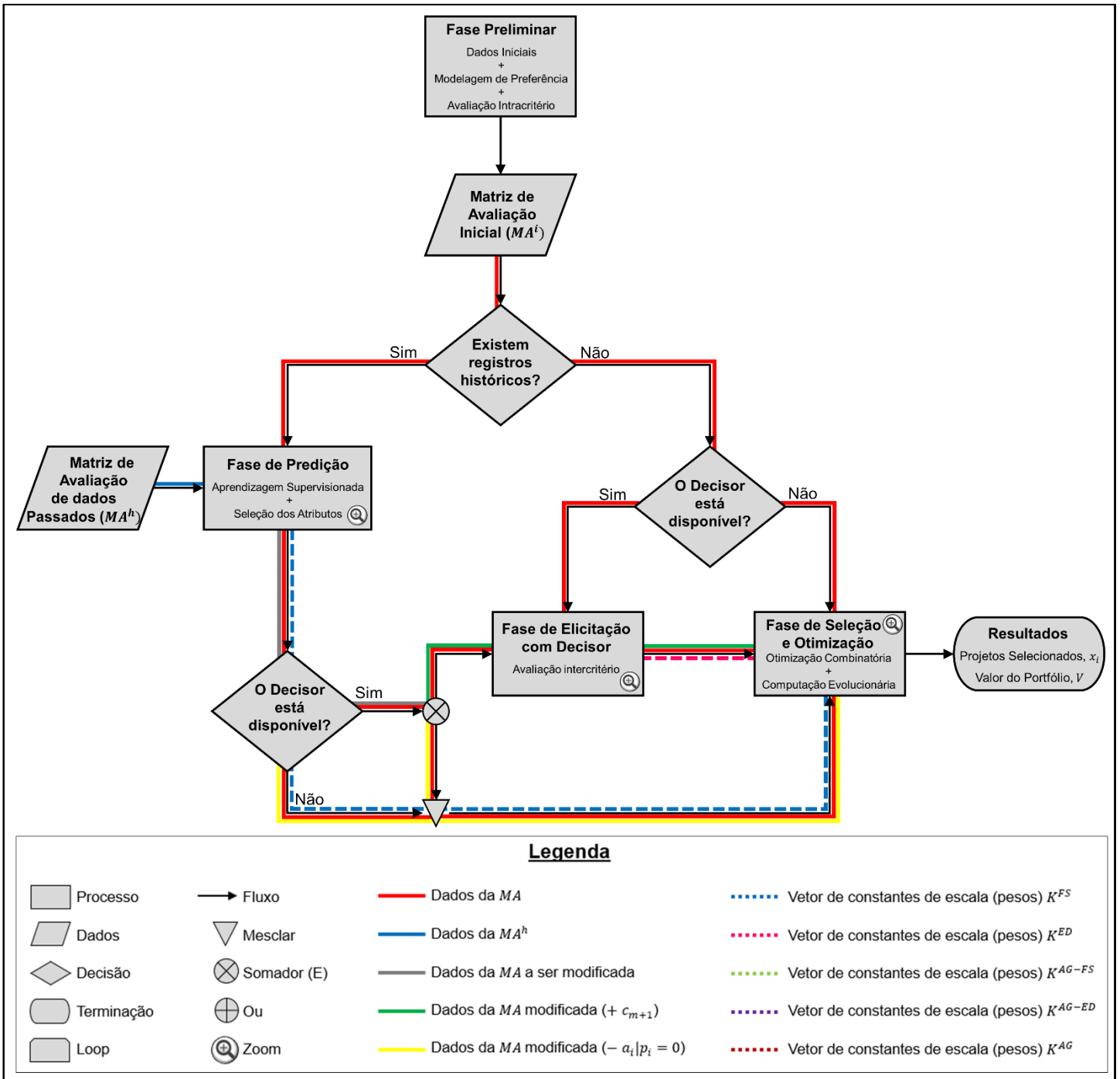


Figura 4.1 - Framework geral do modelo proposto

A legenda do *framework* desta Figura 4.1 informa de maneira clara se uma determinada etapa consiste em um processo, uma decisão, um loop, e assim por diante. Alguns dos itens/ícones inseridos na legenda não aparecem neste *framework* geral para evitar excesso de informações, porém estão presentes nas figuras da descrição de cada uma das fases (seções 5.4 até 5.6) e, para ter-se uma padronização da legenda, optou-se por utilizar a mesma legenda com todos os itens/ícones para todas as figuras. O ícone de “zoom”, que aparece em algumas das etapas, simboliza que o(s) procedimento(s) que compõe(m) aquele determinado

processo ou loop exige(m) maiores detalhes de informações, sendo que estes serão apresentados à parte, em outro fluxograma (estrutura de pseudocódigo). Juntamente com as setas de fluxo do processo de análise, em algumas partes estão presentes linhas coloridas (contínuas para matrizes de avaliações e pontilhadas para vetores de constantes de escala ou pesos) que representam qual(is) informação(ões) acompanham um determinado caminho.

Observa-se que, ao sair da Fase de Predição, por exemplo, onde é possível se ter uma predição da classificação para as futuras alternativas (instâncias/projetos) como “sucesso” ou “insucesso” (resultado este que corresponde ao conjunto \mathbb{Y}), a matriz de avaliação inicial, MA^i , poderá ser modificada (linha contínua cinza) de duas formas possíveis, conforme a resposta da questão “O decisor está disponível?”:

a) se for “Sim”, a matriz será modificada com a introdução de um novo critério ($+c_{m+1}$) no conjunto \mathbb{C} , referente ao resultado da classificação prevista, ou seja, as n alternativas (projetos/instâncias) podem ter neste novo critério (c_{m+1}) um desempenho $d_{i,m+1} = \{0: "insucesso", 1: "sucesso"\}$, como mostra o Quadro 4.3;

b) se for “Não”, a matriz será modificada com a exclusão das alternativas (instâncias/projetos) cuja predição resultou em “insucesso” ($- a_i | y_i = 0$), sem a introdução de um novo critério (Quadro 4.4). Essa mudança é útil quando existe um número grande de alternativas a serem avaliadas, de modo que a exclusão de alternativas “suspeitas” de serem um “insucesso” servirá como um pré-filtro para a análise.

	c_1	c_2	...	c_m	c_{m+1}	$b:$
a_1	d_{11}	d_{12}	...	d_{1m}	$d_{1,m+1}$	b_1
a_2	d_{21}	d_{22}	...	d_{2m}	$d_{2,m+1}$	b_2
...
a_n	d_{n1}	d_{n2}	...	d_{nm}	$d_{n,m+1}$	b_n
						B

Quadro 4.3 - Esquema da Matriz de Avaliação modificada ($+c_{m+1}$)

	c_1	c_2	...	c_m	$b:$	$y:$
a_1	d_{11}	d_{12}	...	d_{1m}	b_1	y_1
a_2	d_{21}	d_{22}	...	d_{2m}	b_2	y_2
...
a_n	d_{n1}	d_{n2}	...	d_{nm}	b_n	y_n
						B

Quadro 4.4 - Esquema da Matriz de Avaliação modificada ($- a_i | y_i = 0$)

Estas duas matrizes modificadas, MA modificada ($+ c_{m+1}$) e MA modificada ($- a_i | y_i = 0$), que correspondem às linhas contínuas verde e amarela, respectivamente, da Figura 4.1, juntamente com a matriz de avaliação inicial MA , representada pela linha contínua vermelha, e todos os cinco vetores de constante de escala (pesos), são as informações que

podem chegar à última fase, de Seleção e Otimização, onde serão determinadas as possíveis combinações para obtenção do resultado.

Por exemplo, se em um problema de decisão multicritério para seleção de portfólio existirem os registros históricos de decisões semelhantes, que tenham um parecer do resultado (“sucesso” ou “insucesso”) da realização de projetos passados, e se for caracterizado um único decisor, e este estiver disponível para eventuais entrevistas para o procedimento de elicitação de forma consistente, o fluxo dos processos de análise do modelo proposto será: Fase Preliminar → “Existem Registos Históricos?”; Se a resposta for “Sim”, então → Fase de Predição → “O decisor está disponível?”; Se a resposta for “Sim”, então → Fase de Elicitação com decisor E mesclagem com o caminho da resposta “Não” → Fase de Seleção e Otimização → Resultados. Esta sequência do fluxo entre as fases, por onde percorrem as informações das matrizes de avaliação e dos vetores de constantes de escala (pesos), é exemplificada na Figura 4.2.

Em um processo de decisão para um problema de seleção de portfólio tem-se inicialmente a construção da matriz de avaliação inicial MA , a partir da Fase Preliminar, e as informações desta matriz, representadas pela linha vermelha contínua, percorre o fluxo do *framework* até a pergunta “Existem Registos Históricos?”, como mostra a Figura 4.2 (a).

Como a resposta é “Sim”, ou seja, existem registros históricos, as informações da matriz de registros históricos, MA^h , representada pela linha contínua azul, juntamente com as informações da matriz de avaliação inicial (linha contínua vermelha) serão utilizadas na Fase de Predição, como mostra a Figura 4.2 (b).

Como resultado da Fase de Predição, Figura 4.2 (c), tem-se a determinação do vetor K^{FS} (linha pontilhada azul), e a modificação da matriz de avaliação inicial (linha contínua cinza), que poderá ser feita seja com a remoção das alternativas cuja predição foi de “insucesso” (MA modificada $(- a_i | y_i = 0)$, linha contínua amarela), ou seja com a introdução de um novo critério “predição” (MA modificada $(+ c_{m+1})$, linha contínua verde). Estas informações serão utilizadas na Fase de Seleção e Otimização.

Com a disponibilidade do decisor no processo decisório, tanto a matriz de avaliação inicial, quanto a matriz modificada $(+ c_{m+1})$, entram na Fase de Elicitação, tendo como resultado a determinação do vetor K^{ED} , cuja informação também será utilizada na Fase de Seleção e Otimização, como mostra a Figura 4.2 (d).

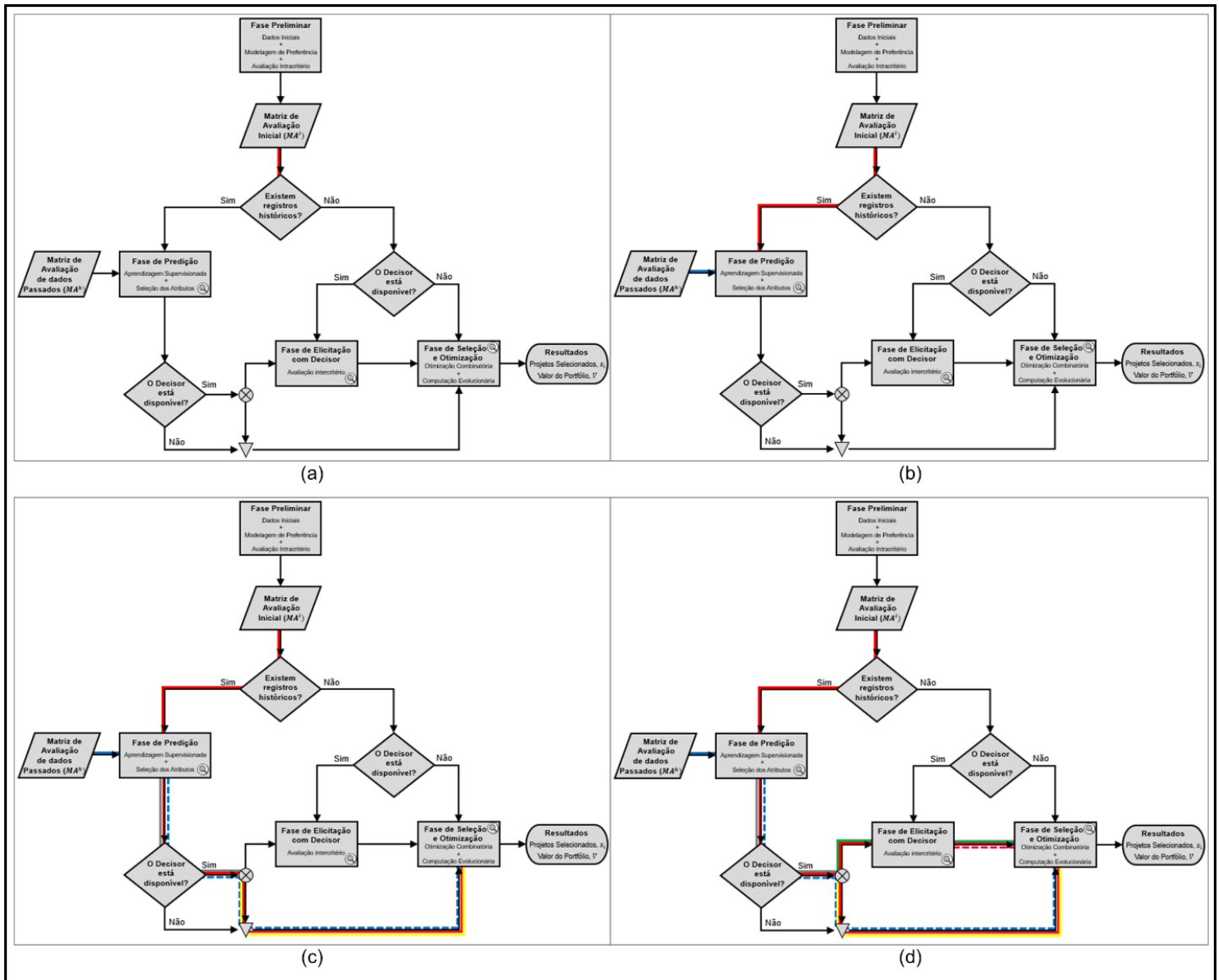





















Figura 4.2 - Sequência do fluxo de informações para o caso de existir registros históricos e o decisor estar disponível

Na Fase de Seleção e Otimização, são determinados os vetores otimizados pelo algoritmo genético (K^{AG} ; K^{AG-FS} e K^{AG-ED}) e, posteriormente, cada uma das três matrizes (MA ; MA modificada ($- a_i | y_i = 0$); ou MA modificada ($+ c_{m+1}$)) será combinada com um dos cinco vetores (K^{FS} ; K^{ED} ; K^{AG} ; K^{AG-FS} e K^{AG-ED}), para a obtenção de um resultado para o problema de seleção de portfólio. Ou seja, para cada resultado do problema, que consiste em determinar os projetos selecionados e valor do portfólio, serão combinados uma única matriz com um único vetor de pesos.

Este caso de existir registros históricos e o decisor estar disponível (“Existem Registros Históricos”? Sim + “O decisor está disponível?” Sim) é o mais completo de todos os casos, em que todos os tipos de matrizes de avaliação e todos os vetores de constantes de escala

(pesos) estarão envolvidos durante o processo. É também o caso mais complexo com relação às possíveis combinações das matrizes de avaliação e vetores de constante de escala (pesos) para obtenção do resultado, número este que pode chegar a 13, como mostra a Tabela 4.2.

Tabela 4.2 - Possíveis combinações para obtenção do resultado

Possíveis Decisões		Matrizes de Avaliações	Vetores de Constantes de escala (peso)	Número de possíveis combinações
Existem Registros Históricos?	O decisor está disponível?			
Sim	Sim	 MA  MA modificada (+ c_{m+1})  MA modificada ($- a_i y_i = 0$)	 K^{FS}  K^{ED}  K^{AG-FS}  K^{AG-ED}  K^{AG}	13
	Não	 MA  MA modificada ($- a_i y_i = 0$)	 K^{FS}  K^{AG-FS}  K^{AG}	6
Não	Sim	 MA	 K^{ED}  K^{AG-ED}  K^{AG}	3
	Não	 MA	 K^{AG}	1

A Tabela 4.2 fornece três informações primordiais: i) quais são os possíveis tipos de matrizes de avaliação que podem chegar à fase de Seleção e Otimização; ii) quais são os possíveis vetores de constante de escala que estarão presentes na referida fase e; iii) o número total das possíveis combinações entre as matrizes de avaliação e vetores de constante de escala (pesos) para a obtenção de um resultado, a depender do fluxo a ser seguido.

Observe-se que, no caso descrito anteriormente (“Existem Registros históricos”? Sim + “O decisor está disponível?” Sim), o número de combinações possíveis entre matrizes de avaliação e vetores de constante de escala (pesos) é de somente 13, e não 15 (número de combinações de 3 matrizes e 5 vetores), pois a matriz de avaliação modificada com a introdução de mais um critério, MA modificada (+ c_{m+1}), só pode ser combinada com os vetores K^{AG} , K^{AG-ED} e K^{ED} , uma vez que não é possível obter um valor para o novo critério (c_{m+1}) nos vetores K^{FS} e K^{AG-FS} , obtidos a partir do procedimento de seleção dos atributos.

O caso mais simples de um problema de decisão multicritério para seleção de portfólio, que este modelo contempla, é aquele no qual não existem registros históricos, nem mesmo a disponibilidade do decisor (“Existem Registros históricos”? Não + “O decisor está disponível?” Não). Neste caso, o fluxo do processo decisório é exemplificado pela Figura 4.3.

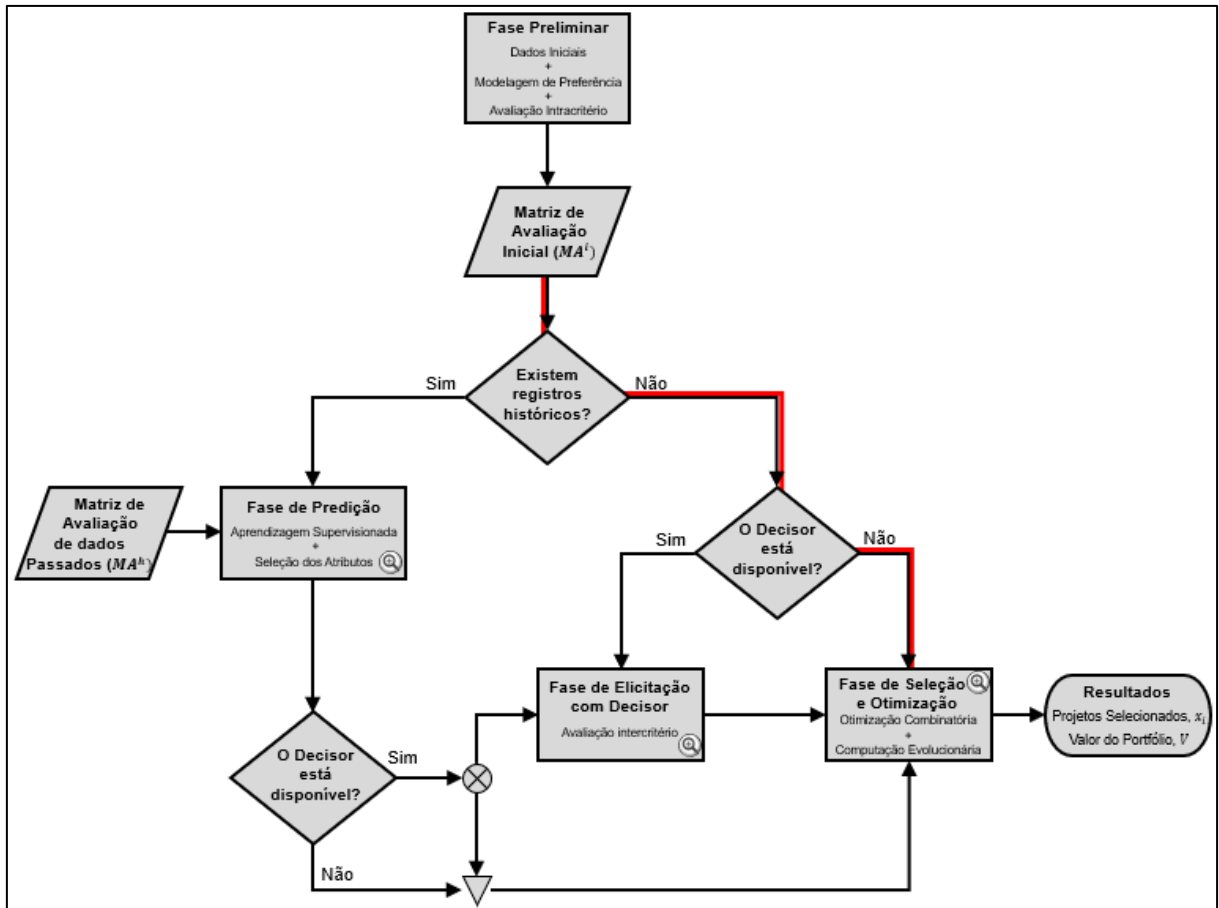


Figura 4.3 - Sequência do fluxo no processo decisório para o caso em que não há registros históricos nem a disponibilidade do decisor

Nesta configuração, mostrada na Figura 4.3, a única informação disponível consiste na Matriz de Avaliação inicial (MA) e no vetor de constante de escala (peso) K^{AG} , gerado de forma aleatória, e sem considerar nenhuma ordem de preferência/importância para a otimização por meio do algoritmo genético. E, como existe uma única forma possível de combinação entre MA e K^{AG} , portanto um único resultado será obtido ao final do processo.

Ressalta-se que, mesmo este sendo um caso possível de análise, e que fornecerá um resultado, em uma situação real de decisão tal resultado deve ser avaliado minuciosamente antes de qualquer implementação da ação, ou conjunto de ações indicadas, uma vez que o grau de confiabilidade do modelo recai sobre uma decisão cujo vetor K^{AG} foi gerado inicialmente de forma aleatória, e sem considerar as preferências de algum decisor ou então sem “aprender” com os casos passados.

E, a elaboração de uma recomendação dos resultados em um processo de análise deve incorporar, não somente o relato do resultado em si, mas também o grau de precisão associado aos parâmetros e dados utilizados, ou seja, deve-se informar os riscos envolvidos no processo. Para o modelo proposto, o grau de confiabilidade está diretamente relacionado com a

quantidade e qualidade das informações (registros históricos e/ou decisor) disponíveis: quanto mais registros históricos, com qualidade, e quanto melhor for a consistência do decisor em fornecer suas preferências, maior será o grau de confiabilidade do modelo.

Sendo assim, recomenda-se que a análise de um problema de decisão multicritério para seleção de portfólio, que utilize o modelo proposto, tenha sempre pelo menos registros históricos ou ao menos um decisor disponível.

4.3. DESCRIÇÃO DA FASE PRELIMINAR

Na Fase Preliminar são estruturados os elementos básicos para a formulação do problema de decisão. É nesta fase que os dados iniciais são definidos, juntamente com a modelagem de preferências e a avaliação intracritério.

Os dados iniciais do procedimento de análise consistem em: a) caracterizar o decisor de forma clara, além de outros atores envolvidos no processo decisório; b) definir os critérios ou atributos que representem os objetivos da decisão para o processo de modelagem quantitativa, ou seja, para aqueles critérios (atributos) que não são “naturais” deve-se estabelecer uma escala de avaliação, que permita medir um conjunto de fatores relacionados ao critério; c) estabelecer o espaço de ações, correspondente a um conjunto discreto de alternativas (instâncias/projetos); d) determinar o conjunto de desempenho de cada alternativa em cada critério; e) determinar o custo para o desenvolvimento de cada alternativa; f) estipular o orçamento total disponível (*budget*); g) verificar se existem registros históricos. Sendo assim, ao determinar os dados iniciais, têm-se condições de construir as matrizes de avaliação inicial, *MA*.

Na modelagem de preferência são definidos: a) a estrutura de preferência do decisor; e b) a racionalidade. Então, primeiramente, deve-se avaliar qual estrutura de preferência é mais adequada para representar as preferências do decisor, e isso afeta diretamente na escolha do modelo e métodos a serem utilizados. Por exemplo, se a estrutura “preferência e indiferença” (P,I) for apropriada para que o decisor possa expressar suas preferências, então pode ser utilizado o modelo tradicional, que inclui métodos de critério único de síntese. Já se o decisor precisar de outras relações de preferências como, por exemplo, a incomparabilidade, deve-se escolher algum método que permita essa condição. E, depois, deve-se determinar a racionalidade mais adequada para o decisor no problema considerado, podendo ser compensatória ou não compensatória, e isso também implica na correta escolha do método a

ser utilizado. Para o modelo proposto, pressupõe-se que a estrutura de preferências do decisor seja do tipo (P, I), e a racionalidade seja compensatória, podendo assim utilizar-se de modelos de agregação aditivo, no qual incluem os métodos de critério único de síntese.

A avaliação intracritério, considerando o pressuposto acima, consiste no estabelecimento da função de valor para cada critério (atributo), $v_j(a_i)$. No caso deste ser uma função linear, será possível realizar um procedimento de normalização “divisão pela soma” diretamente sobre as consequências determinísticas. Porém, enfatiza-se sobre a correta maneira de efetuar o processo de normalização, que deve estar vinculado com a parametrização, e alerta que, para modelo de agregação aditivo em portfólio é necessário o uso de escala de razão. Além disso, se no problema existir algum critério (atributo) que não seja linear, cuidados adicionais deverão ser tomados para representar de forma adequada a função $v_j(a_i)$.

4.4. DESCRIÇÃO DA FASE DE PREDIÇÃO

Caso existam registros históricos de decisões semelhantes, com uma possível avaliação do resultado da escolha de um determinado projeto, ou seja, se uma decisão já tiver sido tomada anteriormente para a seleção de projetos, e tais projetos tenham sido avaliados como, por exemplo, “sucesso” ou “insucesso”, é possível a construção da matriz de dados históricos, MA^h , e dar sequência no processo realizando-se os procedimentos da Fase de Predição.

Essa fase consiste nos procedimentos de Aprendizagem Supervisionada e Seleção dos Atributos (Figura 4.2). O primeiro procedimento diz respeito ao reconhecimento de padrões para predição e classificação de instâncias, utilizando-se de técnicas de aprendizagem de máquina (*machine learning*, seção 2.2). Como consequência, é possível ter uma classificação prevista das alternativas (projetos) como “sucesso” ou “insucesso”, ou seja, a definição do conjunto \mathbb{Y} . Já o segundo procedimento, de seleção dos atributos (*feature selection*, seção 2.2.7), está relacionado ao uso das técnicas de abordagem “*filter*”, para estabelecer um *ranking* dos critérios (atributos). Como resultado, é possível estabelecer o vetor K^{FS} de constantes de escala (pesos) para os critérios.

No fluxograma da Figura 4.4 estão descritos alguns dos métodos e/ou técnicas que podem ser utilizadas em cada uma das etapas, dos procedimentos acima mencionados. Porém,

ressalta-se que estes métodos e/ou técnicas descritos são apenas algumas sugestões, existindo outras possibilidades, e que não necessariamente todos devem ser adotados, para ter-se um resultado satisfatório. Por exemplo, para a etapa de tratamento de dados, não é obrigatório realizar o balanceamento das classes, embora esta seja uma estratégia aconselhável para aumento do desempenho preditivo do modelo, a depender da estrutura dos dados. Ou então, na etapa de treinamento e teste do algoritmo classificador, não é necessário utilizar-se de todos os algoritmos, ou restringir-se somente aos algoritmos mencionados, uma vez que existem diversos outros na literatura.

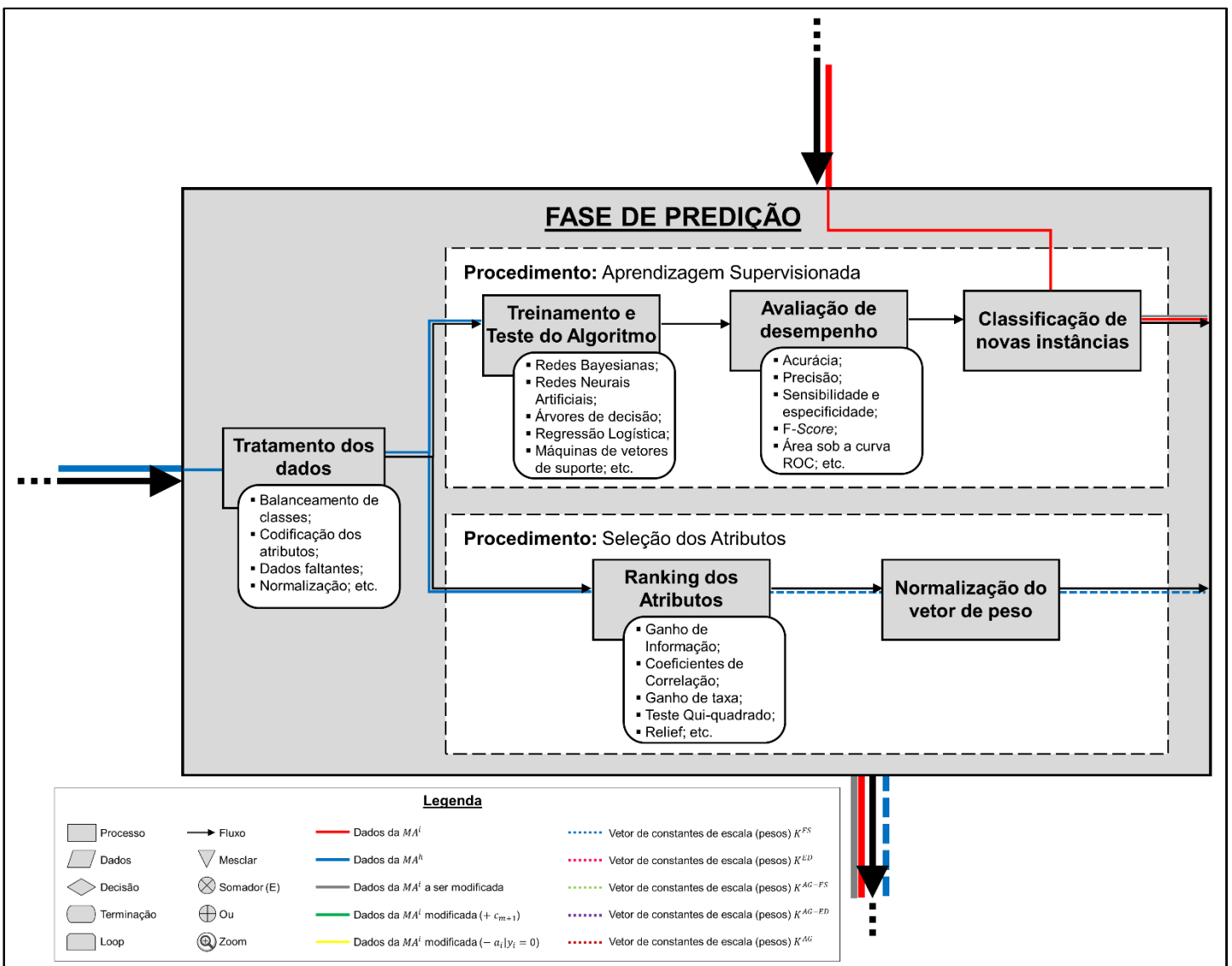


Figura 4.4 - Detalhes dos Procedimentos da Fase de Predição

Em resumo, o fluxo do processo na Fase de Predição inicia-se com as informações da matriz de dados históricos, MA^h (linha contínua azul), que passam por uma etapa de

Tratamento dos dados, cujas técnicas a serem utilizadas dependem de como estão apresentados os dados. No modelo proposto, é fortemente aconselhável que seja feito, pelo menos, o balanceamento de classes. Essas informações tratadas seguem por dois caminhos:

- i. Procedimento de Aprendizagem Supervisionada: onde serão treinados e testados os algoritmos classificadores (seções 2.2.1 a 2.2.5), e aquele que obtiver melhor desempenho na tarefa de predição, medido por alguma das métricas de avaliação (seção 2.2.6), será utilizado para classificação das alternativas (linha contínua vermelha). O resultado desse procedimento (linha contínua cinza) pode ser tratado sob duas formas diferentes, como já mencionado anteriormente, seja com a introdução de um novo critério ($+ c_{m+1}$), ou com a eliminação de alternativas cuja classificação prevista foi de “insucesso” ($- a_i | y_i = 0$).
- ii. Procedimento de Seleção dos atributos: onde serão utilizadas técnicas de *feature selection* com abordagem “*filter*”, para estabelecer um *ranking* dos critérios (atributos) com base no ganho de informação (*information gain*, seção 2.2.7). Como resultado, é possível estabelecer um vetor de constantes de escala (pesos) para os critérios, K^{FS} (linha pontilhada azul), que passará por uma normalização para que a soma dos k_j^{FS} seja igual a um.

4.5. DESCRIÇÃO DA FASE DE ELICITAÇÃO COM DECISOR

É nesta fase que o vetor de constante de escala (pesos) dos critérios, K^{ED} , pode ser definido por meio de um procedimento de elicitação com o decisor. É uma fase que diz respeito à condução do processo de decisão multicritério de uma forma tradicional, ou seja, com a figura de um decisor, cujas preferências serão levadas em conta para a obtenção dos resultados, por meio de um modelo de agregação aditivo determinístico. Em função disso, o procedimento que compõe essa fase é denominado de Avaliação Intercritério (Figura 4.5), que tem como etapa principal a obtenção das constantes de escala k_j^{ED} .

Várias são as maneiras pelas quais é possível definir o vetor de constante de escala (pesos) K^{ED} , ou seja, os métodos para parametrização do modelo aditivo, a citar: os procedimentos de elicitação baseado em *trade-offs*; os procedimentos de *Swing*, referentes aos métodos SMARTS e/ou SMARTER (*Simple Multi-attribute rating technique*; Edwards e Barron, 1994); método *Even Swaps* (Almeida, 2013; Hammond, *et al.*, 1999); método

Macbeth (*Measuring Attractiveness by a Categorical Based Evaluation Technique*; Bana e Costa et al., 2005); método AHP (*Analytic Hierachy Process*; Saaty, 1980), dentre outros.

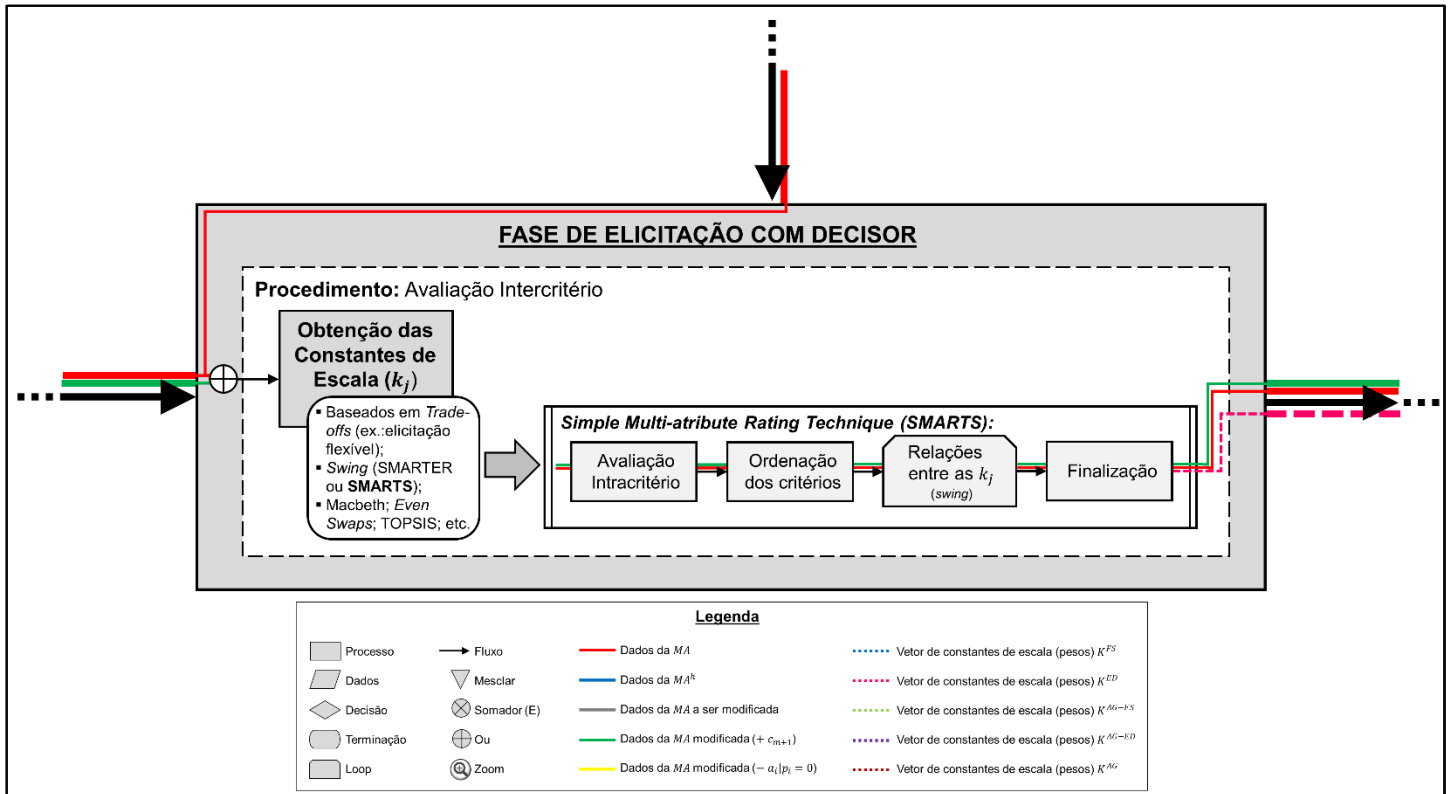


Figura 4.5 - Detalhes dos procedimentos da Fase de Elicitação com Decisor

Para o modelo proposto, como pode ser visto na Figura 4.5, optou-se pelo método SMARTS (*Simple Multi-attribute rating technique*), descrito na seção 2.1. Este método usa o procedimento de pesos por *swing*, e considera funções de valor lineares para a avaliação intracritério, simplificando as hipóteses no processo de análise, o que leva a menores erros de eliciação com o decisor.

As duas etapas iniciais deste método consistem na avaliação intracritério e na ordenação dos critérios, utilizando-se da matriz de avaliação inicial MA (linha contínua vermelha) ou da matriz de avaliação modificada (+ c_{m+1} ; linha contínua verde). Na próxima etapa, para obter as relações entre as constantes de escala k_j , inicia-se um processo interativo de perguntas e respostas com o decisor. E, por fim, tem-se a finalização do método, cuja saída resulta no vetor de constante de escala (peso) K^{ED} (linha pontilhada magenta).

4.6. DESCRIÇÃO DA FASE DE SELEÇÃO E OTIMIZAÇÃO

Esta é a última fase do modelo de análise, na qual uma matriz de avaliação será combinada com um vetor de constante de escala (peso) para gerar um resultado final, que consiste basicamente na resposta de qual(is) projeto(s) foi/foram selecionados para compor o portfólio ($x_i = 1$, se o projeto i foi selecionado; $x_i = 0$, caso contrário), e na mensuração do valor do portfólio (V), obtido pela formulação apresentada em (2.1) e (2.2).

Ressalta-se que será fornecido um resultado (x_i, V) distinto para cada vetor de constante de escala ($K^{FS}; K^{ED}; K^{AG-FS}; K^{AG-ED}; K^{AG}$) combinado com cada matriz de avaliação ($MA^i; MA^i$ modificada $+c_{m+1}; MA^i$ modificada $-a_i | y_i = 0$), obtidos no decorrer do processo de análise a depender da direção do fluxo, como já explicado na seção 4.2 (ver Tabela 4.2).

Os procedimentos (Figura 4.6) que compõem essa fase são:

- i. Otimização Combinatória: consiste na aplicação de algum método para solução do problema de seleção de portfólio que, assim como é especificado no problema da mochila (*Knapsack Problem*), pode ser obtida, dentre outras formas, pela resolução do modelo matemático ou pela enumeração exaustiva, que fornecem a solução ótima, ou então por métodos aproximativos baseados nas abordagens de programação dinâmica ou de *branch-and-bound* (B&B).
- ii. Computação Evolucionária: consiste em um procedimento de otimização meta-heurístico para determinar um vetor de constante de escala (peso) que resulte em um portfólio com maior valor, V , podendo ser utilizado métodos como: algoritmo genético; colônia de formigas; enxame de partículas, entre outros.

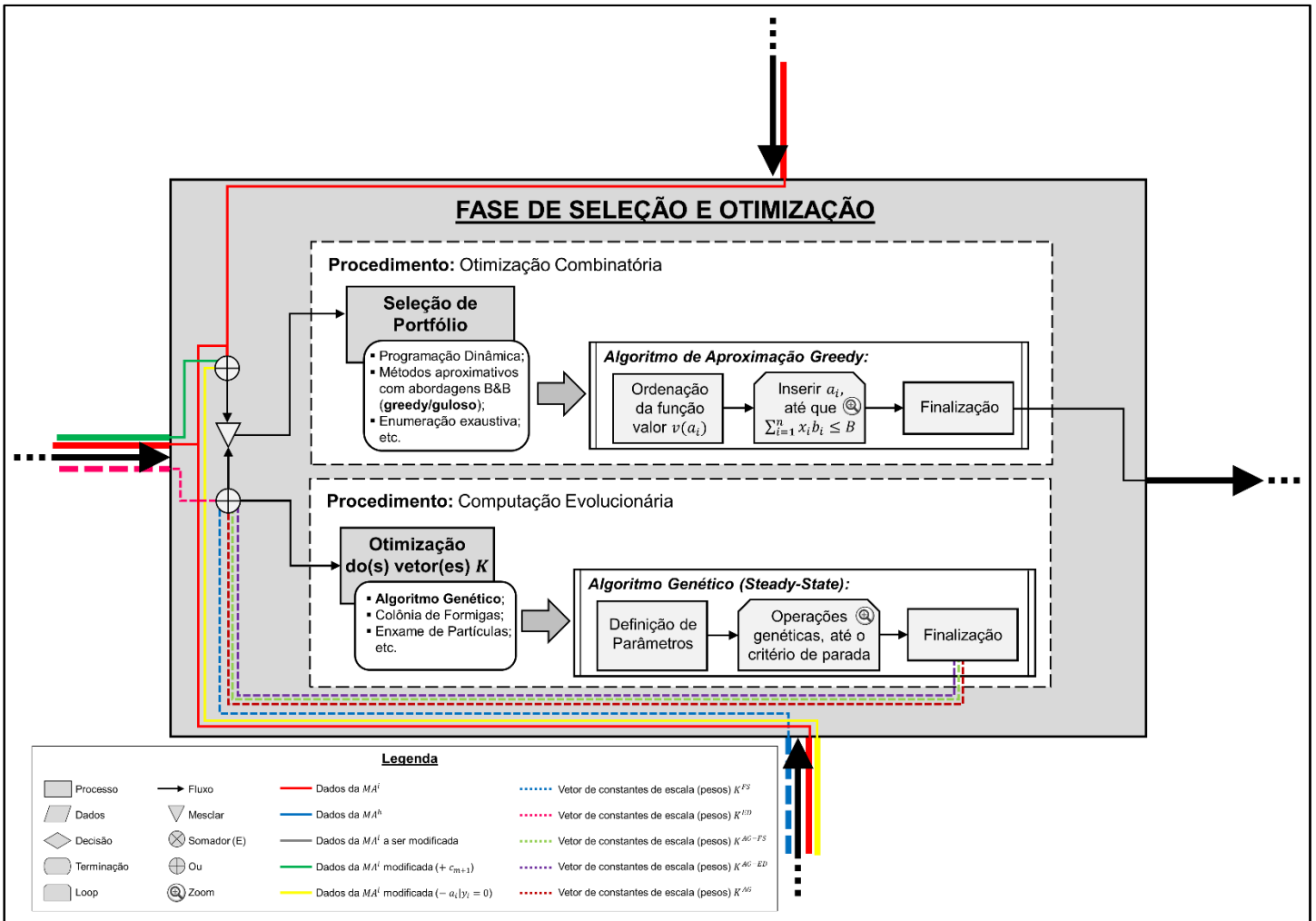


Figura 4.6 - Detalhes dos procedimentos da Fase de Seleção e Otimização

Então, como pode ser visto na Figura 4.6, as informações das matrizes de avaliação (linhas contínuas vermelhas, verde e/ou amarela) e/ou dos vetores de constante de escala/peso (linhas pontilhadas azul e magenta), a depender do fluxo de onde vieram, entram na Fase de Seleção e Otimização, onde será feita a mesclagem de uma única matriz com um único vetor para o procedimento de otimização combinatória e assim, conseqüentemente, obter um resultado.

No modelo proposto, o procedimento de otimização combinatória será feito utilizando-se do Algoritmo de Aproximação *Greedy*, uma vez que nos problemas com muitas alternativas, a resolução por métodos exatos é muito custosa computacionalmente. As etapas deste algoritmo consistem basicamente em: a) ordenar as alternativas pela função valor $v(a_i)$, obtida pela equação (2.3); b) inserir as alternativas no portfólio, até que a restrição, descrita por (2.2), seja satisfeita; e c) finalizar o procedimento. Estas etapas são melhores explicadas pelo pseudocódigo do Quadro 4.5.

Algoritmo 4.1: Algoritmo Greedy para seleção do portfólio

\mathbb{A} e \mathbb{X} definidos na seção 4.1
 $v(a_i) \leftarrow$ função valor da alternativa a_i (calculado pela equação 2.3)
 $B \leftarrow$ orçamento total disponível (budget)
 $b_i \leftarrow$ custo de cada alternativa

```

1  Ordenar  $\mathbb{A}$  conforme  $v(a_i)$ ;
2   $somaP \leftarrow 0$ ;  $V \leftarrow 0$ ;
3   $\mathbb{X} \leftarrow \{0, \dots, 0\}$ ;
4   $i \leftarrow 1$ ;
5  while  $somaP < B$  do
6     $Portfólio \leftarrow a_i$ ; (alternativa  $a_i$  de  $\mathbb{A}$  após ordenação)
7     $x_i \leftarrow 1$ ;
8     $somaP \leftarrow somaP + x_i * b_i$ ;
9     $V \leftarrow V + x_i * v(a_i)$ ;
10 end
11 Retornar  $V$  e  $\mathbb{X}$ 

```

Quadro 4.5 - Pseudocódigo do Algoritmo de Aproximação *Greedy* para seleção do portfólio

No algoritmo de Aproximação *Greedy*, adaptado para seleção de portfólio em um contexto multicritério, as alternativas (projetos) que possuem melhores valores da função $v(a_i)$ serão inseridas primeiramente no portfólio, respeitando a restrição orçamentária do custo de cada alternativa (b_i) e do total disponível (B). A finalização do algoritmo se dá com o retorno do valor do portfólio (V), e da informação das alternativas que compõem (ou não) o portfólio, correspondente ao conjunto \mathbb{X} .

Observe que, na Figura 4.6, os vetores de constante de escala (peso) K^{FS} ou K^{ED} , linhas pontilhadas azul ou magenta, respectivamente, podem também ser inseridos no procedimento de computação evolucionária, onde passam por um processo de otimização, resultando nos vetores denominados como K^{AG-FS} ou K^{AG-ED} , linhas pontilhadas verde e roxo, respectivamente. Nestes casos, os vetores resultantes do procedimento de otimização respeitam a ordem de preferência estabelecida pelo decisor, no caso do vetor K^{ED} ou a ordem de importância estabelecida pelo *feature selection*, no caso do vetor K^{FS} . Mas também, neste procedimento, pode ser criado um vetor de constante de escala (pesos), denominado K^{AG} (linha pontilhada vermelha), independentemente dos vetores criados anteriormente, ou seja, sem considerar a ordem de importância/preferência dos critérios estabelecida pelo decisor ou pelo *feature selection*.

No modelo proposto, o processo de otimização dos vetores de constantes de escala (pesos) será feito utilizando o Algoritmo Genético do tipo *Steady-State* (seção 2.3.1), cujas etapas consistem basicamente em: a) definição dos parâmetros do algoritmo genético; b) aplicação dos operadores genéticos, até que o critério de parada seja satisfeito; e c) finalização do procedimento. Estas etapas são melhores explicadas pelo pseudocódigo do Quadro 4.6.

Algoritmo 4.2: Algoritmo Genético *Steady-State* para otimização do vetor K

```

 $n \leftarrow$  tamanho do cromossomo (número de critérios/atributos)
 $m \leftarrow$  tamanho da população
 $cParada \leftarrow$  critério de parada (número de iterações; máximo fitness; estagnação do fitness)
 $tCross \leftarrow$  taxa de crossover
 $tMut \leftarrow$  taxa de mutação

1  $P \leftarrow \{p_1, p_2, \dots, p_m\}$ ; (população inicial de  $m$  cromossomos de tamanho  $n$ )
2 for  $p_i \in P$ , calcular fitness( $p_i$ );
3 Ordenar  $P$  conforme melhor fitness( $p_i$ );
4 while  $cParada$  não satisfeito do
5    $D \leftarrow$  selectFunction( $P$ ); (cromossomos selecionados para operadores genéticos)
6   if  $\text{rand}(0,1) \leq tCross$  then
7      $D' \leftarrow$  crossover( $D$ );
8     for  $d'_i \in D'$  do
9       if  $\text{rand}(0,1) \leq tMut$  then  $D' \leftarrow D' \cup$  mutação( $d'_i$ );
10     $D' \leftarrow$  reparação( $D'$ );
11  end
12  for  $d'_i \in D'$ , calcular fitness( $d'_i$ )
13   $P \leftarrow$  selecionar  $m$  melhores cromossomos de  $P \cup D'$ ;
14 end
15 Retornar  $p_i \in P$  com melhor fitness

```

Quadro 4.6 - Pseudocódigo do Algoritmo Genético *Steady-State* para otimização dos vetores de constante de escala (pesos)

No Algoritmo Genético, adaptado para a otimização do vetor de constante de escala (peso) do problema de seleção de portfólio, um cromossomo corresponde a um vetor K , e os genes deste cromossomo representam os k_j . Sendo assim, sempre um cromossomo, seja ele da população inicial ou da geração de filhos, deverá ser normalizado para que a soma dos k_j seja igual a “1”.

Uma questão importante a ser notada é que, para o cálculo da função de aptidão ou adequação (*fitness*(p_i)) de cada cromossomo, é necessário executar o procedimento de otimização combinatória (Algoritmo *Greedy*), para obter o resultado do valor do portfólio, V , de cada cromossomo. Ou seja, os cromossomos são ordenados pelo *fitness*, que corresponde a função de valor do portfólio V .

As demais funções presentes no pseudocódigo correspondem a: a) *selectFunction*(): função para selecionar dois indivíduos (cromossomos) para proceder com os operadores genéticos, podendo ser feita pelo método da roleta ou então pela função “*Select*”; b) *crossover*(): operação genética de recombinação, para o cruzamento entre pares de cromossomos, podendo ser de um ou dois pontos; c) *mutação*(): operação genética para a modificação de um ou mais genes de um cromossomo; d) *reparação*(): operação genética para manter a população factível.

Outra observação importante é que, para os vetores de constante de escala (peso) K^{AG-FS} ou K^{AG-ED} , cuja ordem de preferência/importância deve ser mantida, a operação de reparação pode ser necessária, se caso nas operações de *crossover* e/ou mutação tenha a geração de um filho que não respeite a ordem de preferência/importância. Então, neste caso, a restrição está em manter a ordem de preferência/importância, e um cromossomo inactível é aquele que não respeita essa ordem e que, portanto, deve ser reparado para tornar-se factível.

E a etapa de finalização do algoritmo genético consiste no retorno do cromossomo (vetor de constante de escala/peso, K) que obteve maior *fitness*, ou seja, maior valor do portfólio V .

5. IMPLEMENTAÇÃO DA FASE DE SELEÇÃO E OTIMIZAÇÃO

Os procedimentos da Fase de Predição contam com aplicativos/*softwares* que incorporam as técnicas e/ou métodos utilizados, tal como o WEKA (*Waikato Environment for Knowledge Analysis*), desenvolvido e atualizado por pesquisadores da Universidade de Waikato, que possui uma coleção de algoritmos de aprendizado de máquina para tarefas de mineração de dados. E, para os métodos descritos na Fase de Elicitação com decisor, também já existem aplicativos/*softwares* específicos que podem ser utilizados, ou então, a depender do método selecionado e da complexidade do problema, não é necessário o uso de recurso computacional, podendo o cálculo ser feito com auxílio de uma calculadora, por exemplo.

Porém, para a aplicação do modelo proposto em casos reais, foi necessário o desenvolvimento e implantação da Fase de Seleção e Otimização em um *software* específico. Embora já existam aplicativos/*softwares* para ambos os procedimentos desta fase, nenhum é feito de forma integrada e com as especificidades requeridas pelo modelo proposto.

Sendo assim, foram implementados os Algoritmos 4.1 e 4.2, descritos nos Quadros 4.5 e 4.6, respectivamente, em um *software* com uma interface gráfica para os usuários. Entende-se por usuários, neste caso, os especialistas/analistas de um processo decisório, uma vez que são necessários entendimentos sobre as técnicas/métodos para utilizar-se do *software* desenvolvido. Ressalta-se que a implementação destes algoritmos foi feita com a linguagem de programação e interface de desenvolvimento Delphi, sendo que o Algoritmo 4.1 (que consiste no método de seleção de portfólio) foi implementado em parceria com os pesquisadores da Universidade Federal de Pernambuco, e o Algoritmo 4.2 (referente ao AG *Steady-State*) foi desenvolvido pela autora.

A interface gráfica do primeiro procedimento, de Otimização Combinatória, é mostrada na Figura 5.1. Nela pode-se observar campos para a introdução da Matriz de Avaliação (*Consequence Matrix*); do vetor de pesos (*Weights*); do custo equivalente a cada alternativa (*Constraints*); e do orçamento total disponível (*Allowed Cost*). Estes dados devem ser incluídos manualmente na própria interface, ou então pela importação de um arquivo (*Import Worksheet*) no formato .xls (Anexo A)

Ao clicar em “*Start*”, o problema será resolvido a partir do Algoritmo 4.1, mostrando-se os seguintes resultados: a) *Value*: valor de cada portfólio, $v(a_i)$; b) *Output Summary*: solução final de projetos selecionados, custo (*Best Cost*) e valor do portfólio (*Best Value FO*), para os pesos fornecidos na entrada de dados, sem ainda realizar a otimização dos mesmos.

Para dar início ao procedimento de Computação Evolucionária, com o objetivo de otimizar a solução final pela mudança no vetor de pesos, clica-se no botão “*Genetic Algorithm*”, que levará a uma segunda janela do *software*, com a interface gráfica mostrada na Figura 5.2.

Figura 5.1 - Interface gráfica do procedimento de Otimização Combinatória

Neste momento, é necessário incluir os parâmetros do Algoritmo Genético, que são: a) tamanho do cromossomo (*Chromosome Size*), dado pelo número de critérios/atributos do problema; b) Manter ou não uma sequência pré-definida da ordem de preferência/importância dos critérios/atributos (*Keep the sequence* e *Inform the sequence*); c) tamanho da população inicial (*Population Size*), e se deseja incluir um ou mais cromossomo(s) pré-definido(s) nesta população (*Introduce a chromosome*); e) definir a Função de seleção (*Select Function* ou *Roulette Wheel*); f) selecionar os operadores genéticos e suas respectivas taxas (*Simple*

crossover e/ou *Uniform Mutation*), bem como o número de cortes para o *crossover* (*Crossover points*); g) escolher o(s) critério(s) de parada, e definir os respectivos parâmetros (*Number of Generation* e/ou *Convergence Critéria* e/ou *Improvement in the best solution*); h) definir se a função objetivo é de maximização ou minimização (*Evaluation Function*).

Genetic Algorithm applied to MCDM Portfolio Selection Show more Information

Chromosome Representation

Chromosome Size:

Keep the gene sequence
 Yes No

Inform the sequence:

Allowed only integers. Use space to separate numbers.

Initial Population

Population Size:

Introduce a chromosome:

Use comma for decimal number, and space to separate numbers

Selection Function

Select Function Roulette Wheel

Genetic Operators

Simple Crossover - Rate:

Uniform Mutation - Rate:

Crossover Points
 1 2

Termination Criteria

Number of generations:

Convergence Criteria - Threshold:

Improvement in the best solution:

Evaluation Function

Maximization Minimization

Output File Name:

Output Summary:

Figura 5.2 - Interface gráfica do procedimento de Computação Evolucionária

Similarmente, ao clicar em “*Start*”, o problema será resolvido a partir do Algoritmo 4.2, mostrando-se os resultados no campo “*Output Summary*”, que consistem nos cinco melhores cromossomos (vetores de constante de escala/peso) obtidos pela otimização, bem como a criação de um arquivo .txt que mostra a evolução de todas as iterações (Anexo B). Ao finalizar, será mostrado na janela anterior, da Figura 5.1 (*Solution – Weights found by the Genetic Algorithm*), a solução dos projetos selecionados e valor do portfólio, definida a partir do melhor cromossomo encontrado pelo algoritmo genético.

Ambas as interfaces apresentam o botão “*Show more information*”, que contém uma breve explicação de cada procedimento, com a descrição do significado de todos os parâmetros utilizados, como mostra a Figura 5.3.

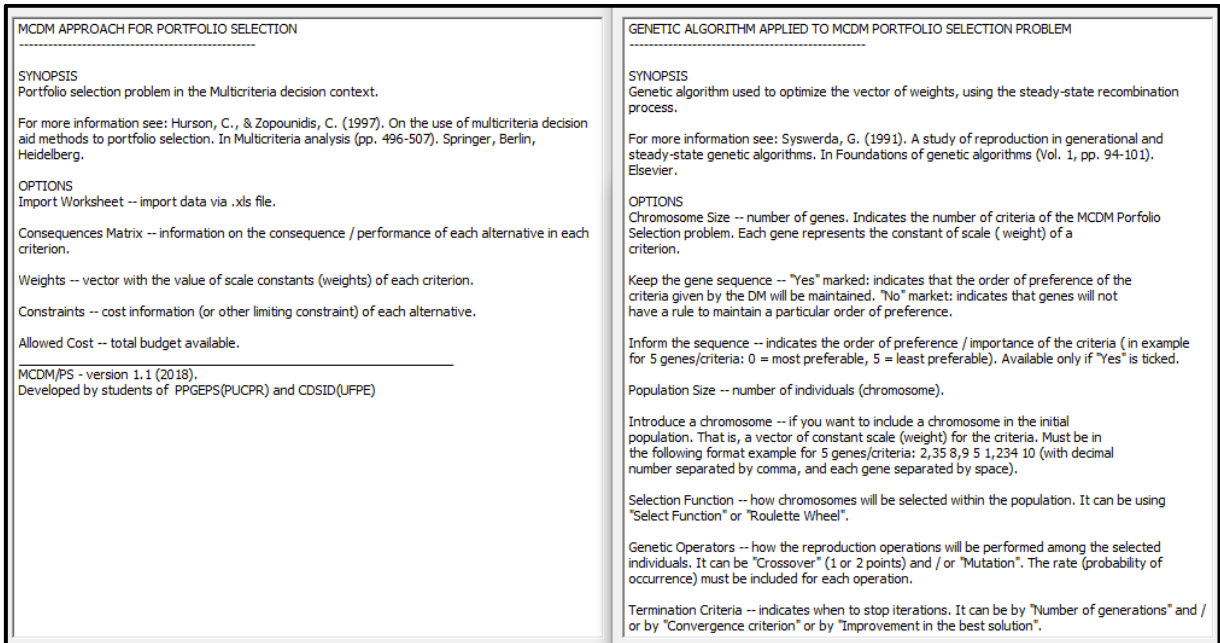


Figura 5.3 - Interface gráfica das informações mostradas com os botões “*Show more information*”

5.1. DEMONSTRAÇÃO DE UM PROBLEMA HIPOTÉTICO

Para uma breve exemplificação do uso e funcionamento do *software* elaborado para a implementação da Fase de Seleção e Otimização, considere um problema de decisão multicritério para seleção de portfólio, com três critérios e cinco alternativas (projetos), cujo objetivo é determinar o portfólio ótimo, de modo a maximizar o valor V dado por (5.1), sujeito a (5.2).

$$\max V = x_1 v(a_1) + x_2 v(a_2) + x_3 v(a_3) + x_4 v(a_4) + x_5 v(a_5) \quad (5.1)$$

$$\text{s. a: } x_1 b_1 + x_2 b_2 + x_3 b_3 + x_4 b_4 + x_5 b_5 \leq B \quad (5.2)$$

em que x_i é a variável de decisão, inteira e binária, de incluir ($x_i = 1$) ou não ($x_i = 0$) a alternativa i no portfólio; b_i corresponde ao custo de realização de cada alternativa i ; B é o orçamento total disponível; e $v(a_i)$ é a função de valor aditiva, obtida por (5.3).

$$v(a_i) = k_1 v_1(d_{i,1}) + k_2 v_2(d_{i,2}) + k_3 v_3(d_{i,3}) \quad (5.3)$$

Na otimização feita pelo procedimento que utiliza o AG, o objetivo é determinar os valores das constantes de escala (k_1 , k_2 e k_3), que resultem em um máximo V .

Então, para o teste deste exemplo hipotético, os dados iniciais do problema foram gerados de forma aleatória, que correspondem ao: desempenho de cada alternativa em cada critério, $d_{i,j}$; custo de cada alternativa, b_i ; orçamento total, B , como mostra a Matriz de Avaliação inicial do Quadro 5.1. O vetor de constante de escala foi definido inicialmente como $K = [0,3; 0,5; 0,2]$.

	c_1	c_2	c_3	b :
a_1	1	2	7	120
a_2	3	4	4	210
a_3	2	3	2	100
a_4	4	7	9	230
a_5	6	9	4	150
				$B = 560$

Quadro 5.1 - Matriz de Avaliação inicial (MA) para o problema hipotético

Estes dados foram introduzidos no *software*, seguindo-se os passos descritos anteriormente para importação e início do procedimento, e ao final obteve-se a solução gerada a partir do vetor de constante de escala inicial, K^{ED} , como mostra a Figura 5.4.

The screenshot shows the 'MCDM approach for Portfolio Selection' software interface. It includes several sections:

- Consequences Matrix:** A table with 5 alternatives (A1-A5) and 3 criteria (X1, X2, X3). The values are: A1 (1, 2, 7), A2 (3, 4, 4), A3 (2, 3, 2), A4 (4, 7, 9), A5 (6, 9, 4).
- Weights:** A table with 3 criteria (X1, X2, X3) and their weights (0,3, 0,5, 0,2).
- Constraints:** A table with 5 alternatives (a1-a5) and their costs (120, 210, 100, 230, 150). The allowed cost is 560.
- Value:** A table with 5 items (1-5) and their values (2,610, 5,110, 3,810, 8,940, 11,400).
- Output Summary:** Displays the solution found by the Genetic Algorithm. The best cost is 500 and the best value FO is 20,8268219451511.

Figura 5.4 - Interface com o resultado obtido para o vetor K^{ED} do problema hipotético

Percebe-se que a solução dada pelo vetor de constantes de escala (pesos) inicial K^{ED} foi a inclusão dos projetos “ a_1 ”, “ a_4 ” e “ a_5 ”, ou seja, $x_1 = x_4 = x_5 = 1$, ao passo que $x_2 = x_3 = 0$. Com estes projetos selecionados, o valor de portfólio (*Best Value FO*) resultou em $V = 20,826$, tendo um custo total (*Best Cost*) de 500 ($120 + 230 + 150$), que não ultrapassa o limite estabelecido de $B = 560$.

Para o procedimento de otimização, com uso do algoritmo genético, os parâmetros foram introduzidos como mostra a Figura 5.5. O tamanho de cromossomos (*Chromossome Size*), que representa o número de critérios, foi de “3”; a ordem de importância dos critérios foi mantida a mesma do vetor K^{ED} , ou seja, com c_2 e c_3 sendo os critérios mais e menos importante, respectivamente; o tamanho da população foi estabelecida em 50, sem a introdução de nenhum cromossomo; a função de seleção escolhida foi a “*Select Function*”; e as configurações para as operações genéticas foram de realizar o crossover de um ponto a uma taxa de 80%, e mutação de 10%; já o critério de parada foi somente pelo número de gerações, estipulado em 1000 iterações.

Genetic Algorithm applied to MCDM Portfolio Selection [Show more Information](#)

Chromosome Representation

Chromosome Size:

Keep the gene sequence: Yes No

Inform the sequence:

Allowed only integers. Use space to separate numbers.

Initial Population

Population Size:

Introduce a chromosome:

Use comma for decimal number, and space to separate numbers

Selection Function

Select Function Roulette Wheel

Genetic Operators

Simple Crossover - Rate:

Uniform Mutation - Rate:

Crossover Points: 1 2

Termination Criteria

Number of generations:

Convergence Criteria - Threshold:

Improvement in the best solution:

Evaluation Function

Maximization Minimization

Output File Name:

Output Summary:

RUN INFORMATION

Chromosome Size: 3
 Keep the gene sequence: Yes (1 0 2)
 Population Size: 50
 Manually added chromosome:
 Selection Function: Select Function
 Genetic Operators:
 Simple Crossover - Rate: 0,8
 Binary Mutation - Rate: 0,1
 Crossover Points: 1
 Termination Criteria:
 Number of generation: 1000
 Evaluation Function: Maximization

SUMMARY: Top five solutions of weights found

0,023 0,964 0,011 - Fitness:24,1557558525555
 0,029 0,946 0,024 - Fitness:24,0325409680581
 0,029 0,946 0,024 - Fitness:24,0325409680581
 0,029 0,946 0,024 - Fitness:24,0325409680581
 0,029 0,946 0,024 - Fitness:24,0325409680581

SUMMARY: Projects selected with the best weight found

3 4 5

Figura 5.5 - Interface com o resultado da otimização pelo algoritmo genético

Como a função de avaliação do algoritmo genético é de maximização, os cromossomos gerados na população inicial e em cada iteração são ordenados do melhor para o pior *fitness* (valor do portfólio, V). Pelo arquivo de saída do procedimento, é possível verificar a evolução dos cromossomos desde a criação da população inicial até a última iteração, como é mostrado parcialmente na Figura 5.6.

GENETIC ALGORITHM OUTPUT	Output of each Iteration
<pre> ----- Run Information ----- Chromosome Size: 3 Keep the gene sequence: Yes (1 0 2) Population Size: 50 Manually added chromosome: Selection Function: Select Function Genetic Operators: Simple Crossover - Rate: 0,8 Binary Mutation - Rate: 0,1 Crossover Points: 1 Termination Criteria: Number of generation: 1000 Evaluation Function: Maximization Initial Population (normalized): Chromosome and Fitness 0,3020 0,5630 0,1330 - Fitness:20,8383885767826 0,3250 0,4380 0,2350 - Fitness:20,6726818418454 [...] 0,3630 0,3790 0,2560 - Fitness:20,3595808305445 Initial Population (normalized and ordered): Chromosome and Fitness 0,0230 0,9640 0,0110 - Fitness:24,1557558525555 0,1710 0,7890 0,0390 - Fitness:22,6048253884727 [...] 0,4830 0,4830 0,0330 - Fitness:19,5769016377236 </pre>	<pre> ----- First Iteration ----- Selected Parents: 0,2630 0,6840 0,0520 0,1830 0,6330 0,1830 Created children: Chromosome and Fitness Value 0,2430 0,5860 0,1690 - Fitness:21,2886416225818 0,1990 0,7440 0,0560 - Fitness:22,2468047351004 New Population (normalized and ordered): Chromosome and Fitness 0,0230 0,9640 0,0110 - Fitness:24,1557558525555 0,1710 0,7890 0,0390 - Fitness:22,6048253884727 0,2140 0,7570 0,0280 - Fitness:22,2482449650869 0,1990 0,7440 0,0560 - Fitness:22,2468047351004 [...] 0,4270 0,4520 0,1200 - Fitness:19,6746255330519 ----- [...] continuing the iterations ----- Last Iteration ----- Selected Parents: 0,0290 0,9460 0,0240 0,0290 0,9460 0,0240 Created children: Chromosome and Fitness Value 0,0290 0,9460 0,0240 - Fitness:24,0325409680581 0,0290 0,9460 0,0240 - Fitness:24,0325409680581 New Population (normalized and ordered): Chromosome and Fitness 0,0230 0,9640 0,0110 - Fitness:24,1557558525555 0,0290 0,9460 0,0240 - Fitness:24,0325409680581 [...] 0,0290 0,9460 0,0240 - Fitness:24,0325409680581 </pre>

Figura 5.6 - Evolução de cada iteração do procedimento de otimização pelo algoritmo genético

Neste problema hipotético, a população inicial de 50 indivíduos foi gerada aleatoriamente com três genes em cada cromossomo, sendo normalizada (*Initial Population – normalized*) para que a soma destes seja igual a “1”, e ordenada pelos maiores *fitness*, inicialmente de 24,1557 (*Initial Population – normalized and ordered*).

Na primeira iteração, foram selecionados dois indivíduos (*Selected Parents*), e realizado o *crossover* para a criação dos filhos (*Created children*), que necessitaram de reparação para obedecer a ordem de importância imposta e, posteriormente, de normalização. Nota-se que nesta iteração não houve mutação, pois somente dois filhos foram criados, sendo estes obtidos pelo *crossover*. Também, já nesta iteração o segundo filho criado, com *fitness* de 22,2468, ocupou a quarta posição na nova população (*New Population*). O processo continua até a última iteração, que resultou em um cromossomo com *fitness* máximo de 24,1557.

Este cromossomo representa o vetor de constante de escala (pesos) otimizado a partir do algoritmo genético, e que respeita a ordem de importância/preferência dos critérios pré-estabelecida pelo decisor, ou seja, $K^{AG-ED} = [0,023; 0,964; 0,011]$. A solução fornecida com este vetor resultou na seleção dos projetos “ a_3 ”, “ a_4 ” e “ a_5 ”, tendo o valor do portfólio (*Best Value FO*) igual a 24,1557, e um custo total (*Best Cost*) de 480, como mostra a Figura 5.7.

Output Summary	
Solution (weights given by the Decision Maker):	
1	Best Cost: 500
4	Best Value FO: 20,8268219451515
5	
Solution (weights found by the Genetic Algorithm):	
3	Best Cost: 480
4	Best Value FO: 24,1557558525555
5	

Figura 5.7 - Interface parcial com o resultado das soluções encontradas para o problema hipotético

Sendo assim, o resultado pós-otimização com o vetor K^{AG-ED} , além de selecionar outros projetos para compor o portfólio, proporcionou uma redução no custo (de 500 para 480) e um aumento no valor do portfólio (de 20,8268 para 24,1557) em relação a solução anterior, dada pelo vetor K^{ED} .

6. APLICAÇÕES DO MODELO: ESTUDOS DE CASOS

Para a aplicação e validação do modelo proposto, foram selecionados dois casos reais que envolvem a problemática de seleção de portfólio no contexto de decisão multicritério, e que se diferenciam principalmente pela presença de registros históricos, o que leva a caminhos diferentes no fluxograma do modelo proposto. Os casos são: **Caso I - Seleção de projetos para empresa do setor elétrico** (“Existem Registos Históricos”? Não + “O decisor está disponível?” Sim); **Caso II - Avaliação de risco de crédito para pessoa jurídica** (“Existem Registos Históricos”? Sim + “O decisor está disponível?” Sim).

O primeiro caso diz respeito a uma empresa privada brasileira do setor elétrico, que atua na geração, distribuição e comercialização de energia. A empresa é responsável por controlar 13 distribuidoras, localizadas nos estados de Minas Gerais, Paraíba, Sergipe, Rio de Janeiro, Mato Grosso, Mato Grosso do Sul, Tocantins, Paraná e São Paulo, e está presente em 788 municípios, atendendo a 6 milhões de unidades consumidoras, o que corresponde ao total de 16 milhões de pessoas.

São vários os novos projetos candidatos a compor o portfólio da empresa para serem desenvolvidos, provenientes de diversas áreas, tais como: os projetos exigidos pela agência reguladora e de planejamento de energia elétrica; demanda de tecnologia e informação (TI); projetos de P&D; projetos de planejamento estratégico; e contratos de clientes.

Sendo assim, o problema do Caso I, dito em termos gerais, consiste em definir os projetos que são realmente relevantes, com potencial de agregar valor ao grupo empresarial e que resultam em melhoria do serviço oferecido aos clientes, sendo limitado a uma restrição orçamentária, definida em aproximadamente R\$4 bilhões, valor este que representa o montante disponível pela empresa para a realização de novos projetos. Pesquisas desenvolvidas para seleção de portfólio, e que relacionam a empresa em questão, podem ser encontradas em Bortoluzzi *et. al.* (2017), Martins *et. al.* (2017) e em López e Almeida (2014). Dados parciais para este Caso I encontram-se no Anexo C.

O segundo caso, diferentemente da ideia comum que se têm da seleção de projetos e portfólio, traz uma adaptação do problema da avaliação de risco de crédito para o contexto de projetos e portfólio, em que será analisada a concessão de crédito bancário para pessoa jurídica. Sendo assim, cada operação de crédito solicitada por empresas será tratada como um projeto, e o problema neste caso consiste em definir quais operações serão contempladas para a concessão do crédito, com o objetivo de minimizar o risco de perdas, ou seja, selecionar os projetos que têm maior credibilidade, sujeito a uma ou mais restrições.

Para tanto, deve-se considerar que existem restrições quanto ao limite máximo de exposição por cliente, fixado em 25% (vinte e cinco por cento) do Patrimônio de Referência, estabelecido pela Resolução nº 4.379 do Comitê Monetário Nacional (CMN), de 30 de outubro de 2014 (Banco Central do Brasil, 2014), e além disso, ao limite máximo total de empréstimo concedido pela instituição financeira, que deve respeitar, basicamente: a) o Acordo de Capital da Basileia, oficialmente denominado *International Convergence of Capital Measurement and Capital Standards*, divulgado pelo Comitê de Supervisão Bancária da Basileia (*Basel Committee on Banking Supervision*, BCBS) em 1988, e suas versões revisadas, Basileia II em 2004, e Basileia III em 2010, que têm como “objetivo criar exigências mínimas de capital para instituições financeiras como forma de fazer face ao risco de crédito” (Banco Central do Brasil, 2018); b) limites para operações de crédito, definidos pela Resolução CMN nº 1.559/1998, e suas atualizações (Banco Central do Brasil, 1998); c) limites de operações de crédito para empresas por área de atuação, exigidos pelo governo, e também estipuladas em resoluções do CMN, como por exemplo, os limites de crédito rural.

Como o modelo proposto, a princípio, permite a inclusão de apenas uma restrição, não podendo assim contemplar todas as limitações descritas acima, será estipulado uma restrição orçamentária **hipotética** de R\$ 898.410,65 para a concessão de todos os pedidos de empréstimo por empresas, em um determinado período.

Para a investigação do problema no Caso II, além da presença do decisor, foram considerados os registros históricos de uma grande instituição financeira nacional. Foram coletado dados de mais de 39 mil operações de crédito, realizadas por mais de 5 mil clientes (pessoa jurídica ou empresa), só no Estado do Paraná (PR), no período de janeiro de 1996 a junho de 2017. Dados parciais para este Caso II encontram-se no Anexo D.

Ressalta-se ainda que a aplicação do modelo no Caso II, no qual fez-se uso de dados reais de um banco comercial, foi desenvolvido somente para fins de demonstração do funcionamento do modelo proposto, já que nestas situações, a avaliação de concessão de crédito é realizada de forma individual e contínua a cada solicitação de crédito. O modelo terá maior aplicabilidade se utilizado em bancos de fomento (para os quais os dados não puderam ser obtidos), como exemplo o BNDES (Banco Nacional de Desenvolvimento), em editais de projetos de inovação, recuperação, entre outros tipos de projetos, podendo ser uma ferramenta de grande potencial no apoio a tomada de decisão neste tipo de instituição financeira.

As seções subsequentes, 6.1 e 6.2, descrevem detalhadamente o fluxo do modelo proposto para os problemas dos casos I e II, respectivamente.

6.1. SELEÇÃO DE PROJETOS PARA EMPRESA DO SETOR ELÉTRICO

Como mencionado anteriormente, o Caso I contempla um problema em que o decisor está disponível para o procedimento de elicitação, porém não possui registros históricos de projetos passados. Sendo assim, o problema será resolvido seguindo-se o fluxo mostrado na Figura 6.1. E, de acordo com a Tabela 4.2, as combinações das matrizes de avaliação e vetores de constante de escala (peso) resultarão em três possíveis resultados.

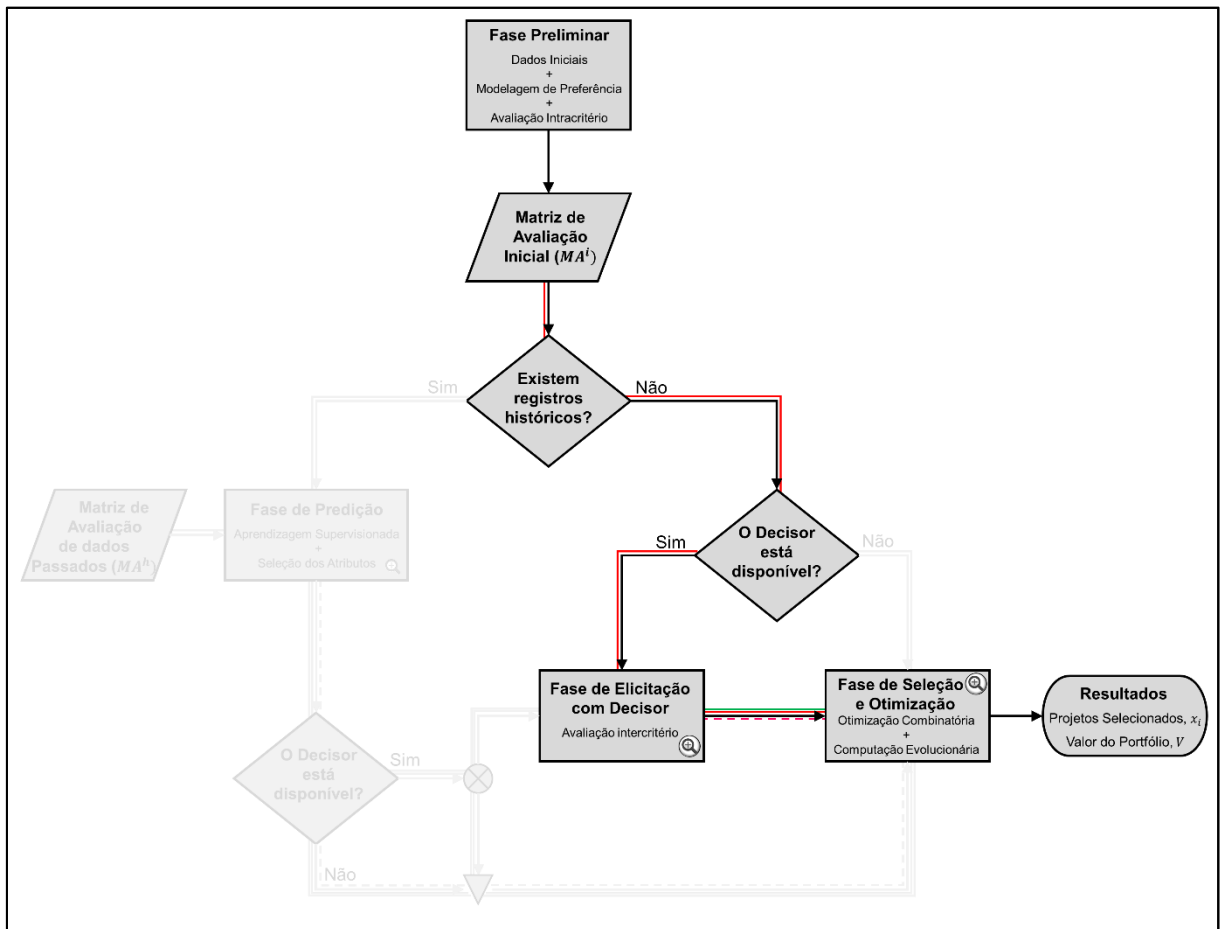


Figura 6.1 - Fluxo do modelo proposto para o Caso I: Seleção de projetos para empresa do setor elétrico

6.1.1. Caso I: Fase Preliminar

Na fase preliminar tem-se a definição dos dados iniciais, a modelagem de preferência e a avaliação intracritério. Para o problema descrito no Caso I, os dados iniciais consistem em:

- a) Atores do Processo decisório:

- Decisor: Diretor da Companhia de Energia Elétrica;
 - Analista(s): a autora.
- b) Conjunto de Critérios (\mathbb{C}): total de seis critérios, escolhidos pelo próprio decisor e alinhados com os objetivos estratégicos e financeiros da empresa, a partir dos parâmetros definidos como relevantes na seleção dos projetos, sendo eles:
- c_1 = Impacto no Resultado;
 - c_2 = Alinhamento Estratégico;
 - c_3 = Melhoria de Indicadores Regulados;
 - c_4 = Contribuição para Satisfação dos Clientes;
 - c_5 = Complexidade;
 - c_6 = Probabilidade de Sucesso.
- c) Conjunto de Alternativas (\mathbb{A}): total de 40 projetos, propostos por diversas áreas da empresa, dos quais: 10 projetos para a Expansão da margem de lucro (AM), 10 projetos para o Risco Regulatório (RR), 10 projetos para a Satisfação do Cliente (SC) e 10 projetos para Negócio Competitivo (NC).
- a_1 = AM01;
 - ...;
 - a_{40} = NC10.
- d) Conjunto de desempenho de cada alternativa em cada critério (\mathbb{D}): definidos subjetivamente pelo decisor, por uma escala verbal definida em “Alto”, “Médio” e “Baixo”, que posteriormente será convertida para uma escala numérica ordinal, para obter-se uma informação quantitativa (Tabela 6.1).
- $d_{1,1}$ = “Médio” (desempenho da a_1 no c_1);
 - ...;
 - $d_{40,6}$ = “Alto” (desempenho da a_{40} no c_6).
- e) Conjunto de custos para o desenvolvimento de cada projeto (\mathbb{B}): informação inerente ao requerimento da inclusão do projeto no processo de seleção da empresa,
- b_1 = R\$ 172.018.877,00;

- ... ;
- $b_{40} = \text{R\$ } 293.764.626,00$.

f) Orçamento total disponível:

- $B = \text{R\$ } 4.101.963.667,00$

g) Inexistência de registros históricos da realização de projetos passados.

Para a modelagem de preferências do decisor, como já mencionado anteriormente, o modelo proposto exige as pressuposições de que a estrutura de preferência seja do tipo (P, I) , e que a racionalidade seja compensatória.

A Tabela 6.1 relaciona o conjunto de critérios estabelecidos, com as escalas verbal e numéricas e suas respectivas descrições dos níveis considerados. Também está descrita nesta tabela a avaliação intracritério, que estabelece uma função de valor $v_j(a_i)$ linear para cada critério, obtida por meio de um procedimento de normalização, em que $v_j(a) = 1$ para a alternativa com valor de desempenho máximo, e $v_j(a) = 0$ para a alternativa com valor de desempenho mínimo, em um determinado critério j .

Tabela 6.1 - Critérios e seus respectivos níveis e função de valor para o Caso I

c_j	Critério	Níveis do Critério		Função de valor $v_j(a_i)$
		Escala Numérica	Descrição (escala verbal)	
c_1	Impacto no Resultado	3	Alto	1
		2	Médio	0,5
		1	Baixo	0
c_2	Alinhamento Estratégico	3	Alto	1
		2	Médio	0,5
		1	Baixo	0
c_3	Melhoria de Indicadores Regulados	3	Alto	1
		2	Médio	0,5
		1	Baixo	0
c_4	Contribuição para Satisfação dos Clientes	3	Alto	1
		2	Médio	0,5
		1	Baixo	0
c_5	Complexidade	3	Alto	0
		2	Médio	0,5
		1	Baixo	1
c_6	Probabilidade de Sucesso	85%	Alto	1
		60%	Médio	0,5
		35%	Baixo	0

Observa-se que, somente no critério c_5 , “Complexidade”, a função de valor é do tipo de minimização, ou seja, quanto menor a complexidade, melhor será o desempenho do projeto neste critério. Para todos os outros critérios, têm-se funções de valor a serem maximizadas.

6.1.2. Caso I: Matriz de Avaliação Inicial

A partir da definição dos dados iniciais, foi possível estruturar a Matriz de Avaliação inicial para o problema proposto no Caso I, que se encontra no Anexo E.

6.1.3. Caso I: Fase de Elicitação com o Decisor

A fase de elicitación com o decisor, realizada para a parametrização do modelo aditivo, foi baseada no procedimento de *trade-off*, ou seja, as constantes de escala foram obtidas em um contexto de elicitación de escala intervalar, devendo-se assim tomar os devidos cuidados (Almeida, Vetschera e Almeida 2014). Diante deste problema de escala, é necessária a conversão apropriada das constantes de escala (pesos) para uma escala de razão.

Sendo assim, o vetor de constante de escala (pesos), obtido por meio do procedimento de elicitación com o decisor, é $K^{ED} = [0,286; 0,228; 0,171; 0,142; 0,114; 0,057]$. Fica então estabelecida a ordem de preferência do decisor, definida por: $c_1 > c_2 > c_3 > c_4 > c_5 > c_6$. No caso, $c_j > c_{j+1}$ indica que o critério c_j é preferível ao critério c_{j+1} .

Vale salientar que a aplicação dessa fase na prática foi desenvolvida por Bortoluzzi *et al.* (2017), em que os autores realizaram a entrevista pessoalmente com o decisor, e aplicaram o método adequado para a definição dos pesos, ou seja, do vetor K^{ED} .

6.1.4. Caso I: Fase de Seleção e Otimização

Nesta fase, a matriz de avaliação inicial, MA , do Anexo E, foi a única submetida aos testes, com os diferentes vetores de constante de escala (pesos). Primeiro, o procedimento de Otimização Combinatória foi feito com essa matriz juntamente com o vetor definido anteriormente, K^{ED} , o que resultou em uma solução.

Dada a ordem de preferência dos critérios por este vetor de constantes de escala (pesos) do decisor, foi realizado o processo de otimização do mesmo, por meio de um procedimento da Computação Evolucionária, mais especificamente, por meio do Algoritmo Genético (AG). Os parâmetros considerados no AG, que serão mantidos sempre os mesmos para todos os testes, foram:

- Tamanho do cromossomo: 6;
- Manter uma sequência pré-definida: Sim, $c_1 > c_2 > c_3 > c_4 > c_5 > c_6$;
- Tamanho da população inicial: 50;
- Incluir um ou mais cromossomo(s): nenhum;
- Função de seleção: *Select Function*;
- Operadores genéticos: *Crossover* (taxa 0.8; e 1 ponto de corte) e *Mutação* (taxa 0.1);
- Critério de parada: Número de gerações/iterações (1000);
- Função objetivo: Maximização.

A partir deste procedimento de otimização, o melhor vetor de constantes de escala (pesos) resultante, em 10 repetições realizadas (Anexo F), foi $K^{AG-ED} = [0,5; 0,5; 0,0; 0,0; 0,0; 0,0]$ que, como pode ser visto, respeita a ordem $c_1 > c_2 > c_3 > c_4 > c_5 > c_6$, se desconsiderado o arredondamento dos valores. Este vetor, combinado com a matriz de avaliação inicial no procedimento de Otimização Combinatória, resultou em uma outra solução para o problema. A solução deste melhor vetor, juntamente com as estatísticas básicas das repetições feitas com o procedimento, estão demonstradas da Tabela 6.2.

Tabela 6.2 - Melhor vetor K^{AG-ED} encontrado, e estatísticas básicas das 10 repetições feitas do AG

	Vetor de Constantes de escala (pesos) K^{AG-ED}						Solução	
	c_1	c_2	c_3	c_4	c_5	c_6	Custo para realização	Valor Portfólio (V)
Melhor vetor	0,500	0,500	0,000	0,000	0,000	0,000	4094251665	86,25
Média	0,4959	0,4952	0,0044	0,0024	0,000	0,000	4096509813	86,13
Desvio Padrão	0,0030	0,0035	0,0039	0,0030	0	0	1190148,49	0,0972

Pela média e desvio padrão de cada um dos valores das constantes de escala (pesos) dos critérios, bem como do valor do portfólio (V), observa-se que não houve muita variação nos resultados a partir das repetições feitas com o procedimento de otimização, sugerindo-se assim a convergência do algoritmo genético em um ponto de ótimo local.

E, por fim, a terceira solução para o problema foi obtida a partir do vetor K^{AG} , criado pelo AG, também em 10 repetições do procedimento (Anexo G), com os mesmos parâmetros definidos anteriormente, porém sem manter a sequência pré-definida da ordem de preferência dos critérios (*Keep the sequence: No*). O melhor vetor resultante foi $K^{AG} = [0,0; 0,069; 0,0; 0,0; 0,0; 0,930]$, como mostra a Tabela 6.3, que tem a ordem $c_6 > c_2 > c_1 \gtrsim c_3 \gtrsim c_4 \gtrsim c_5$. No caso, o símbolo \gtrsim representa a relação de preferência ou equivalência, visto que não foi possível identificar a real ordem de preferência dos critérios, por conta do arredondamento dos valores.

Tabela 6.3 - Melhor vetor K^{AG} encontrado, e estatísticas básicas das 10 repetições feitas do AG

	Vetor de Constantes de escala (pesos) K^{AG}						Solução	
	c_1	c_2	c_3	c_4	c_5	c_6	Custo para realização	Valor Portfólio (V)
Melhor vetor	0,000	0,069	0,000	0,000	0,000	0,930	4059578228	98,48
Média	0,001	0,5306	0,0046	0,0062	0,0009	0,4553	4080698307	97,41
Desvio Padrão	0,0032	0,3629	0,0087	0,0163	0,0028	0,3741	18709431,30	0,7440

Diferentemente dos valores de média e desvio padrão do vetor K^{AG-ED} , discutidos anteriormente, a média e o desvio padrão do vetor K^{AG} , para os valores das constante de escala (pesos) de cada um dos critérios, indicam uma variação entre os critérios c_2 e c_6 durante as repetições, ou seja, em alguns dos resultados c_2 foi o critério mais preferível, enquanto que em outros resultados foi o critério c_6 o mais preferível. Mesmo com essa variação na ordem de preferência dos critérios, o resultado do valor do portfólio (V) manteve-se pouco inalterado, com média de 97,41.

6.1.5. Caso I: Resultados

As três diferentes soluções encontradas para o Caso I, discutidas anteriormente, estão resumidas na Tabela 6.4. A partir da combinação da Matriz de Avaliação Inicial (MA^i) com os três vetores de constante de escala (K^{ED} ; K^{AG-ED} ; K^{AG}), foram obtidas as três soluções para o problema do Caso I. Em cada uma das soluções é possível notar que houve mudanças nos projetos que foram selecionados para compor o portfólio e, conseqüentemente, alterações no custo total para a realização dos projetos.

Tabela 6.4 - Soluções encontradas para o Caso I

Matriz de Avaliação	Vetor de Constantes de escala (pesos)	Ordem de preferência dos critérios	Solução		
			Projetos selecionados	Custo para realização	Valor do portfólio
MA	$K^{ED} = [0,286; 0,228; 0,171; 0,142; 0,114; 0,057]$	$c_1 > c_2 > c_3 > c_4 > c_5 > c_6$	1 3 4 5 6 9 10 11 12 13 15 17 19 20 23 24 26 27 28 30 31 32 33 34 35 36 38 39	4081775825	74,25
	$K^{AG-ED} = [0,5; 0,5; 0,0; 0,0; 0,0; 0,0]$	$c_1 > c_2 > c_3 > c_4 > c_5 > c_6^*$	1 2 3 4 5 6 8 10 12 13 15 17 19 23 24 26 30 31 32 33 34 35 36 38 39 40	4094251665	86,25
	$K^{AG} = [0,0; 0,069; 0,0; 0,0; 0,0; 0,930]$	$c_6 > c_2 > c_1 \succsim c_3 \succsim c_4 \succsim c_5^{**}$	1 4 5 6 10 11 12 13 16 17 19 23 24 25 26 27 28 29 30 32 33 34 35 36 37 38 39	4059578228	98,48

* $c_j > c_{j+1}$ indica que o critério c_j é preferível ao critério c_{j+1} , desconsiderando o arredondamento dos valores;

** $c_j \succsim c_{j+1}$ indica que o critério c_j é preferível ou equivalente ao critério c_{j+1} .

A primeira solução, com o vetor K^{ED} , teve 28 projetos selecionados, com o custo intermediário de R\$ 4081775825, enquanto que a segunda e terceira soluções, com os vetores K^{AG-ED} e K^{AG} , tiveram 26 e 27 projetos selecionados, com o maior e menor custo total, respectivamente. Considerando que todas as soluções estão com a restrição orçamentária dentro do permitido, a primeira solução, que engloba um número maior de projetos atendidos, poderia ser definida como sendo a melhor.

Porém, com relação ao valor do portfólio (V), a primeira solução, com o vetor K^{ED} que representa as constantes de escala obtidas com a elicitación das preferências do decisor, resultou em um menor valor (74,25), quando comparada às soluções com os vetores K^{AG-ED} e K^{AG} , que representam as constantes de escala obtidas pelo procedimento de otimização, cujos valores tiveram um acréscimo de 16,16% e 32,63%, respectivamente. Sendo assim, o resultado da terceira solução, obtida com o vetor K^{AG} , além de representar um menor custo total dentre as soluções, com 27 projetos selecionados, ainda apresentou um melhor valor de portfólio (98,48).

6.2. AVALIAÇÃO DE RISCO DE CRÉDITO PARA PESSOA JURÍDICA

Neste caso de análise, o problema será resolvido com a disponibilidade das informações do decisor e, também, com existência dos registros históricos. Sendo assim, baseando-se no modelo proposto, o fluxo do processo de decisão será como esquematizado na Figura 6.2. E, como já mencionado anteriormente, este é o caso mais completo e complexo de

análise, no qual o número de combinações de matrizes de avaliação com os vetores de constante de escala (peso) será de 13 possíveis resultados (Tabela 4.2, seção 4).

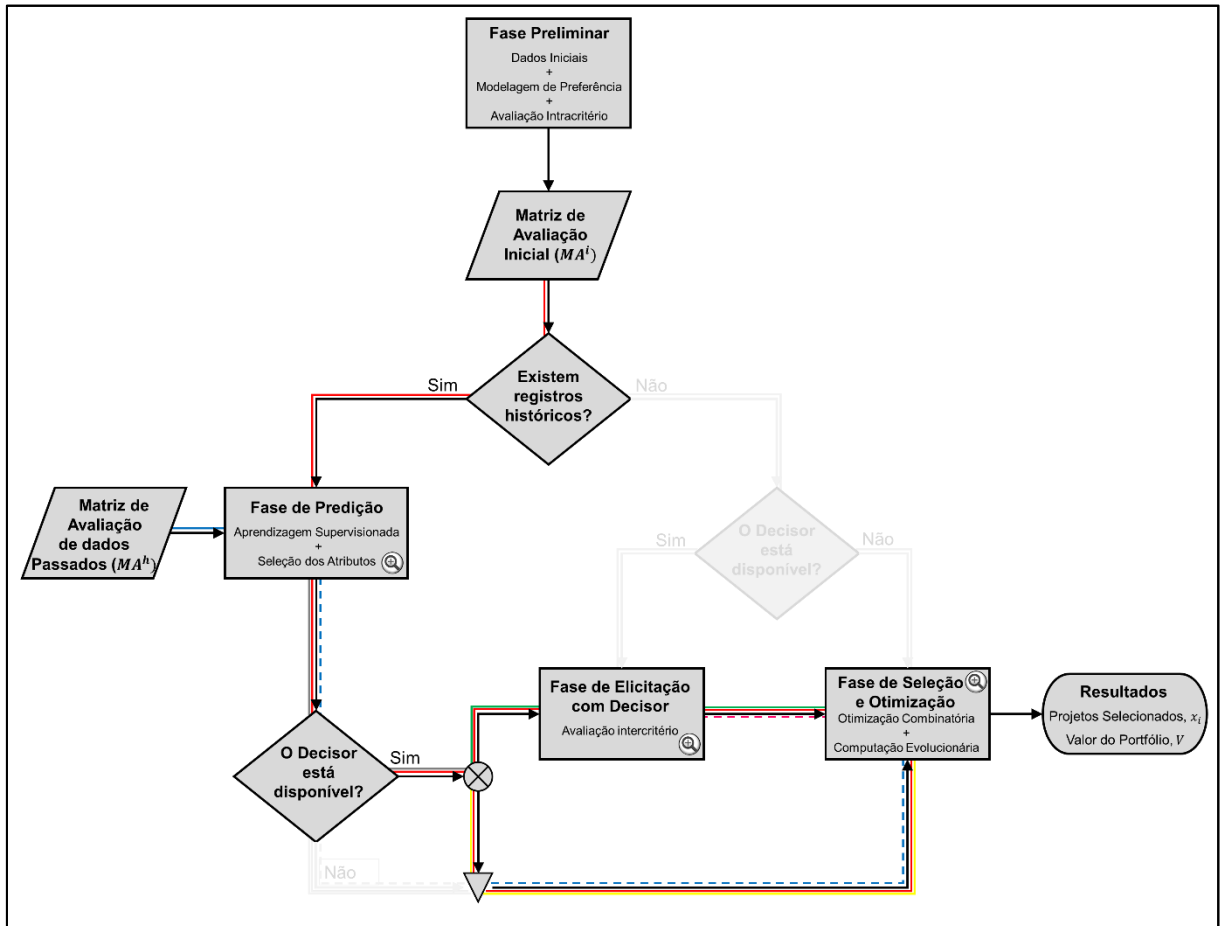


Figura 6.2 - Fluxo do modelo proposto para o Caso II: Avaliação de risco de crédito para pessoa jurídica

6.2.1. Caso II: Fase Preliminar

Para o problema descrito no Caso II, os dados iniciais consistem em:

a) Atores do Processo decisório:

- Decisor: Assessor da Superintendência Regional de Curitiba;
- Cliente: Professor, especialista em problemas bancários;
- Analista(s): a autora.

b) Conjunto de Critérios (C): total de sete critérios, escolhidos pela analista a partir de um banco de dados disponibilizado pelo decisor, sendo eles:

- c_1 = Faturamento bruto anual;
 - c_2 = Quantidade de Funcionários;
 - c_3 = Tempo de Empresa;
 - c_4 = Tempo de Cliente;
 - c_5 = Segmento;
 - c_6 = Risco;
 - c_7 = Restrição.
- c) Conjunto de Alternativas (\mathbb{A}): total de 40 operações de crédito, solicitadas pelas pessoas jurídicas a instituição financeira, durante o mês de junho de 2017
- a_1 = Operação 1 (solicitada pela Empresa 404);
 - ...;
 - a_{40} = Operação 40 (solicitada pela Empresa 5394).
- d) Conjunto de desempenho de cada alternativa em cada critério (\mathbb{D}): informações fornecidas pelas empresas, ou definidas pela instituição financeira, no momento da solicitação de crédito.
- $d_{1,1}$ = “1.133.600” (desempenho da a_1 no c_1);
 - ...;
 - $d_{40,7}$ = “Sem Restrições” (desempenho da a_{40} no c_7).
- e) Conjunto de custos para o desenvolvimento de cada projeto (\mathbb{B}): montante de crédito solicitado em cada operação,
- b_1 = R\$ 5.000,00;
 - ...;
 - b_{40} = R\$ 979,27
- f) Orçamento total disponível:
- B = R\$ 898.410,65
- g) Existência de registros históricos da realização de projetos passados, ou seja, as mais de 39 mil operações realizadas anteriormente pelas 5432 empresas, sendo descritos pelos conjuntos \mathbb{A}^h ; \mathbb{D}^h ; e \mathbb{Y}^h :

- $a_1^h = \text{“Empresa 1”}$; $d_{1,1}^h = \text{“2.512.000”}$, ..., $d_{1,7}^h = \text{“Restrição Impeditiva”}$;
 $y_1^h = \text{“Adimplente”}$
- ...;
- $a_{5432}^h = \text{“Empresa 5432”}$; $d_{5432,1}^h = \text{“54.800”}$, ..., $d_{1,7}^h = \text{“Sem Restrições”}$;
 $y_{5432}^h = \text{“Inadimplente”}$

Uma consideração importante a ser feita quanto ao procedimento adotado para a determinação do conjunto \mathbb{Y}^h neste caso, é que as empresas foram classificadas como “Adimplente” ($y_i^h = 1$: *sucesso*) ou “Inadimplente” ($y_i^h = 0$: *insucesso*).

Para isso, todas as mais de 39 mil operações foram verificadas quanto à existência ou não de atrasos, podendo ter operações: a) sem atrasos no pagamento ou que ainda estão dentro do prazo de pagamento; b) com atraso no pagamento inferior ou igual a 60 dias; ou c) com atraso no pagamento superior a 60 dias. Assim, cada uma das empresas foi mapeada em função do número de operações resultantes de cada tipo (uma empresa pode ter uma ou mais operações de cada tipo), e classificada como y_i^h , conforme a equação (6.1),

$$y_i^h(x_i^{>60}) = \begin{cases} 0: \text{“Inadimplente”} & \text{se } x_i^{>60} > 0 \\ 1: \text{“Adimplente”} & \text{caso contrário} \end{cases} \quad (6.1)$$

onde $x_i^{>60}$ representa o número de operações com atraso no pagamento superior a 60 dias que a empresa i possui. Portanto, se uma empresa i qualquer possuir pelo menos uma operação com atraso no pagamento superior a 60 dias, esta será classificada como “Inadimplente” e, se ela não possuir operações com atraso superior a 60 dias, será classificada como “Adimplente”. Sendo assim, do total de empresas (5432) existentes nos registros históricos, 3896 foram classificadas como “Adimplente”, e 1536 empresas classificadas como “Inadimplente”.

O procedimento de modelagem de preferências do decisor, assim como no Caso I descrito anteriormente, considera a exigência do modelo proposto das pressuposições de que a estrutura de preferência seja do tipo (P, I) , e que a racionalidade seja compensatória.

A Tabela 6.5 relaciona o conjunto de critérios estabelecidos com os respectivos níveis considerados, definidos a partir do banco de dados disponibilizado pelo decisor. As variáveis que eram inicialmente contínuas foram discretizadas, com base nos quartis, e as variáveis categóricas permaneceram como estavam. Após isso, todos foram transcritos por uma escala numérica ordinal. Também está descrita nesta tabela a avaliação intracritério, que estabelece uma função de valor $v_j(a_i)$ linear para cada critério, em que $v_j(a) = 1$ para a alternativa com

valor de desempenho máximo, e $v_j(a) = 0$ para a alternativa com valor de desempenho mínimo, em um determinado critério j .

Tabela 6.5 - Critérios e seus respectivos níveis e função de valor para o Caso II

Critério		Níveis do Critério		Função de valor $v_j(a_i)$
c_j	Descrição	Escala Numérica ordinal	Escala Categórica Discreta	
c_1	Faturamento bruto anual	4	3158540,45 a 769665000	1
		2	1372791,64 a 3156893,44	0,66
		3	536200,1 a 1372514,63	0,33
		1	0 a 536200	0
c_2	Quantidade de Funcionários	4	12 a 5113	1
		3	5 a 11	0,66
		2	2 a 4	0,33
		1	0 a 1	0
c_3	Tempo de Empresa	4	15/02/1939 a 01/06/1997	1
		3	03/06/1997 a 19/02/2004	0,66
		2	20/02/2004 a 19/02/2009	0,33
		1	26/02/2009 a 02/12/2016	0
c_4	Tempo de Cliente	4	23/06/1975 a 29/10/2003	1
		3	30/10/2003 a 26/02/2008	0,66
		2	27/02/2008 a 15/06/2011	0,33
		1	16/06/2011 a 05/12/2016	0
c_5	Segmento	6	Estadual	1
		5	Corporate	0,8
		4	Média empresa	0,6
		3	Empresa	0,4
		2	Pequena empresa	0,2
		1	Microempresa	0
c_6	Risco	5	A	1
		4	B	0,75
		3	C	0,5
		2	D	0,25
		1	E	0
c_7	Restrição	3	Sem restrições	1
		2	Restrição fraca	0,5
		1	Restrição Impeditiva	0

6.2.2. Caso II: Matriz de Avaliação Inicial e de registros históricos

A partir da definição dos dados iniciais, foi possível estruturar a Matriz de Avaliação inicial (MA) para o problema proposto no Caso II, bem como a Matriz de Avaliação de registros históricos (MA^h), que se encontram no Anexo H e I, respectivamente.

6.2.3. Caso II: Fase de Predição

O procedimento de aprendizagem supervisionada conta com o treinamento e teste dos algoritmos classificadores, a avaliação dos mesmos conforme alguma métrica de desempenho, e o uso do algoritmo com o melhor desempenho para a classificação de novas instâncias.

Para isso, foi utilizada a Matriz de Avaliação de registros históricos (MA^h), que contém 5432 instâncias/empresas no total, sendo 3896 classificadas como “Adimplente”, e 1536 como “Inadimplente”. O evidente desbalanceamento ou desproporção do número de instâncias em cada classe pode fazer com que as instâncias da classe minoritária (“Inadimplente”) tenham “fortes” chances de serem classificadas incorretamente. Em função disto, foi aplicada uma técnica (*Random Undersampling*) para solucionar o problema de desbalanceamento, que consiste na eliminação aleatória de instâncias da classe majoritária (“Adimplente”), de tal forma que as duas classes tenham o mesmo número (1536) de instâncias.

Sendo assim, das 5432 instâncias da matriz de registro históricos (MA^h), 3072 instâncias foram treinadas e testadas com cinco algoritmos classificadores: Redes Bayesianas (RB); Regressão Logística (RL); Redes Neurais Artificiais (RNA); Árvores de Decisão (AD); e Máquina de Vetores de Suporte (MVS). Estes algoritmos estão implementados no software WEKA[®] (*Waikato Environment for Knowledge Analysis*) e os testes foram realizados com a opção de validação cruzada (*k-fold*), com $k = 10$ partições e com a configuração padrão em todos os parâmetros.

Os algoritmos considerados foram avaliados conforme as seguintes métricas: acurácia; precisão; *recall*; F-score; Estatística Kappa; e Área sob a curva ROC (*Receiver Operating Characteristic*). A Tabela 6.6 mostra os resultados obtidos, com a comparação das diferenças significativas entre as medidas de desempenho obtidas pelos algoritmos, para a qual foi utilizada o Teste *t* para amostras pareadas, com um nível de significância, $\alpha = 0,05$.

Tabela 6.6 - Resultado das métricas de desempenho dos cinco algoritmos classificadores

Métrica de Avaliação	Algoritmos Classificadores				
	RB	RL	RNA	AD	MVS
Acurácia	83,98	82,93*	83,84	80,24*	82,45*
Precisão	0,86	0,84*	0,86	0,80*	0,84*
Recall	0,82	0,82	0,81	0,80*	0,80*
F-score	0,84	0,83*	0,83*	0,80*	0,82*
Kappa	0,68	0,66*	0,68	0,60*	0,65*
AUC-ROC	0,90	0,89*	0,90	0,87*	0,82*

Os algoritmos com melhores valores em cada métrica de avaliação estão destacados na tabela, com sombreamento em cinza. Para a análise de comparação, a primeira coluna de valores (RB) foi dita como referência para verificar se existe ou não diferença significativa com os demais modelos, sendo que as indicações com “*” e “v”, ao lado dos valores, indicam que existem diferenças significativamente menores e maiores, respectivamente, em relação ao padrão de referência, em uma determinada métrica de avaliação. E, no caso de não existir diferença significativa entre um algoritmo classificador com o padrão de referência, foi apresentado somente o valor de desempenho em si, sem nenhuma indicação.

Por exemplo, o algoritmo RB obteve o melhor valor em acurácia (83,98), porém esse valor não possui uma diferença significativa em relação a acurácia do algoritmo RNA (83,84), diferentemente dos valores obtidos pelos demais algoritmos classificadores, em que é possível afirmar que existe uma diferença significativamente menor (representado pelo “*”), como é o caso do RL que tem acurácia de 82,93.

A escolha de qual algoritmo utilizar para a classificação de novas instâncias foi feita com base no número de vezes que obteve o melhor valor nas métricas de desempenho, juntamente com a relação de diferenças significativas, em que é observada claramente a prevalência do algoritmo RB. Sendo assim, o treino da aprendizagem supervisionada com o algoritmo RB foi aplicado na matriz de avaliação inicial MA , para fazer uma predição das 40 instâncias, sendo que 31 foram classificadas como “Adimplentes” e nove como “Inadimplentes”.

A partir deste resultado de predição, foram elaboradas as duas matrizes de avaliação modificadas: uma com a exclusão de instâncias classificadas como “inadimplentes” ($-a_i | y_i = 0$), e outra com a introdução de um novo critério ($+c_{m+1}$), contendo o resultado da predição. Ambas as matrizes modificadas se encontram nos Anexos J e K, respectivamente.

Na matriz modificada ($-a_i | y_i = 0$) é possível verificar a exclusão das alternativas a_3 , a_{11} , a_{13} , a_{25} , a_{29} , a_{34} , a_{36} , a_{37} e a_{39} , que tiveram uma classificação da predição como “Inadimplente”, e na matriz modificada ($+c_{m+1}$) é introduzido um novo critério c_8 , cujas as alternativas são classificadas em uma escala numérica ordinal, em que “1” representa “Adimplente”, e “0” é equivalente a “Inadimplente”. Estas matrizes são o resultado principal do procedimento de aprendizagem supervisionada, e serão utilizadas ao longo do processo descrito pelo *framework* de análise

Outro procedimento da fase de predição é a seleção dos atributos (*feature selection*), que também foi realizado com uso da Matriz de Avaliação de registros históricos (MA^h), aplicando-se o método de ordenamento (*ranker*) a partir do ganho de informação (*information gain*). O procedimento proporcionou o resultado dos critérios/atributos mais relevantes para a classificação das empresas, sendo este resultado descrito na Tabela 6.7, juntamente com os valores normalizados que compõe o vetor K^{FS} .

Tabela 6.7 - Resultado do *Feature Selection* para composição do vetor de constante de escala, K^{FS}

<i>Ranking</i> (<i>Information Gain</i>)	Descrição do critério	Valores normalizados para composição do Vetor K^{FS}
1° (0,304402)	c_6 : Risco	0,475069
2° (0,292975)	c_7 : Restrição	0,457235
3° (0,013204)	c_4 : Tempo de Cliente	0,020607
4° (0,010940)	c_3 : Tempo de Empresa	0,017074
5° (0,010326)	c_5 : Segmento	0,016115
6° (0,008094)	c_1 : Faturamento Bruto Anual	0,012632
7° (0,000812)	c_2 : Quantidade de Funcionários	0,001267

Os valores normalizados foram obtidos por meio da divisão do respectivo valor de ganho de informação pela soma total (0,640753), de tal modo que a soma dos valores normalizados para a composição do vetor K^{FS} resulte em “1”.

Sendo assim, o vetor de constante de escala (pesos), definido a partir do procedimento de seleção dos atributos, é $K^{FS} = [0,012; 0,001; 0,017; 0,020; 0,016; 0,475; 0,457]$, tendo a ordem de preferência dos critérios descrita por $c_6 > c_7 > c_4 > c_3 > c_5 > c_1 > c_2$. Este vetor, combinado com a Matriz de Avaliação inicial (MA), trará a primeira solução possível para o problema e, quando combinado com a matriz de avaliação modificada ($- a_i | y_i = 0$), será obtida a segunda solução. Lembrando que não é possível combinar este vetor com a matriz de avaliação modificada ($+ c_{m+1}$) para a obtenção de uma solução, uma vez que não se tem um valor para o critério adicional.

6.2.4. Caso II: Fase de Elicitação com o decisor

Na fase de elicitação com o decisor, foi aplicado o procedimento *Swing*, referente ao método SMARTS (*Simple Multi-Attribute Rating Technique*; seção 2.1), para ordenação e determinação dos pesos dos critérios. Este procedimento foi realizado pela analista junto ao ator do processo decisório denominado como cliente (Professor, especialista em assuntos bancários), que representou o decisor durante a elicitação.

Os critérios/atributos selecionados para a análise do caso em questão foram apresentados para o cliente, e então efetuado o procedimento *Swing*, primeiramente para a ordenação e determinação dos pesos dos sete critérios listados na Tabela 7.5. A ordem dos critérios ficou definida como: $c_7 > c_1 > c_6 > c_4 > c_5 > c_2 > c_3$. E, em uma escala de zero a 100, foram atribuídos os pesos para cada um dos critérios, como mostra a Tabela 6.8. Os valores dos pesos de cada critério foram normalizados pelo procedimento de divisão pela soma total (345), de tal modo que a soma dos valores normalizados para a composição do vetor K^{ED} resulte em “1”.

Tabela 6.8 - Resultado da elicitación com o Decisor para composição do vetor de constante de escala, K^{ED}

Ranking (<i>Swing</i> – método SMARTS)	Descrição do critério	Valores normalizados para composição do Vetor K^{ED}
1° (100)	c_7 : Restrição	0,289855
2° (90)	c_1 : Faturamento Bruto Anual	0,260869
3° (85)	c_6 : Risco	0,246376
4° (35)	c_4 : Tempo de Cliente	0,101449
5° (20)	c_5 : Segmento	0,057971
6° (10)	c_2 : Quantidade de Funcionários	0,028985
7° (5)	c_3 : Tempo de Empresa	0,014492

Sendo assim, o vetor de constante de escala (pesos) para os sete critérios, definido a partir do procedimento *Swing* conforme as preferências do cliente/decisor, é representado por $K^{ED} = [0,260; 0,028; 0,014; 0,101; 0,057; 0,246; 0,289]$. Nota-se uma semelhança entre a ordem dos critérios estipulada pelo decisor com a ordem obtida anteriormente pelo procedimento de seleção dos atributos (*feature selection*), que se destaca a predominância dos critérios Restrição (c_7) e Risco (c_6) dentre os três primeiros mais importantes em ambas as ordens.

Uma vez que na fase anterior foi construída a matriz de avaliação modificada, com a introdução de um novo critério ($+c_{m+1}$), contendo o resultado da predição de classificação das novas instâncias como “Adimplente” ou “Inadimplente”, nesta fase de elicitación com o decisor foi necessário definir o peso deste novo critério.

Então, o procedimento de *Swing* foi novamente aplicado, desta vez com oito critérios, sendo que o oitavo critério (c_8 : “Predição”) representa a predição do cliente ser classificado como “Adimplente” ou “Inadimplente”, com base em um algoritmo classificador que tem acurácia de aproximadamente 0,85, ou seja, em 85% das vezes o algoritmo consegue classificar corretamente uma instância.

Segundo o cliente/decisor este critério estaria em quarto lugar no *ranking* geral, tendo a nota de 50 em uma escala de zero a 100, como pode ser visto na Tabela 6.9. Novamente os valores foram normalizados (procedimento de divisão pela soma), para que a soma dos valores que compõe o vetor resulte em “1”.

Tabela 6.9 - Resultado da elicitação com o Decisor para composição do vetor de constante de escala, K^{ED} , com a adição de mais um critério (c_8 : Predição)

<i>Ranking</i> (<i>Swing</i> – método SMARTS)	Descrição do critério	Valores normalizados para composição do Vetor K^{ED}
1° (100)	c_7 : Restrição	0,253164
2° (90)	c_1 : Faturamento Bruto Anual	0,227848
3° (85)	c_6 : Risco	0,215189
4° (50)	c_8 : Predição	0,126582
5° (35)	c_4 : Tempo de Cliente	0,088607
6° (20)	c_5 : Segmento	0,050632
7° (10)	c_2 : Quantidade de Funcionários	0,025316
8° (5)	c_3 : Tempo de Empresa	0,012658

No caso de ter os oito critérios na análise, o vetor de constante de escala (pesos), definido a partir do procedimento *Swing* conforme as preferências do cliente/decisor, possui a ordem $c_7 > c_1 > c_6 > c_8 > c_4 > c_5 > c_2 > c_3$, e é representado por $K^{ED} = [0,227; 0,025; 0,012; 0,088; 0,050; 0,215; 0,253; 0,126]$.

6.2.5. Caso II: Fase de Seleção e Otimização

Na fase de seleção e otimização, as três matrizes de avaliação (inicial e modificadas) foram submetidas aos testes, com os diferentes vetores de constante de escala (pesos) obtidos tanto pelo procedimento de seleção de atributos (K^{FS}), quanto pelo procedimento de elicitação com o decisor (K^{ED}).

Dada a ordem de preferência/importância dos critérios destes vetores de constantes de escala (pesos), foi realizado o procedimento da Computação Evolucionária, mais especificamente, a otimização pelo AG, com os seguintes paramentos:

- Tamanho do cromossomo: 7;
- Manter uma sequência pré-definida: “Sim”, para os vetores K^{AG-FS} e K^{AG-ED} ; e “Não” para os vetores K^{AG} ;
- Tamanho da população inicial: 50;

- Incluir um ou mais cromossomo(s): nenhum;
- Função de seleção: *Select Function*;
- Operadores genéticos: *Crossover* (taxa 0.8; e 1 ponto de corte) e *Mutação* (taxa 0.1);
- Critério de parada: Número de gerações/iterações (1000);
- Função objetivo: Maximização.

Para exemplificar os testes que foram realizados com o AG, considere a primeira linha da Tabela 6.10. Esta linha indica que a Matriz de Avaliação inicial (MA) foi submetida ao teste com o AG, sendo mantida a ordem de importância do vetor K^{FS} , ou seja, $c_6 > c_7 > c_4 > c_3 > c_5 > c_1 > c_2$. Em 10 repetições deste procedimento, foi tomado o melhor vetor, que resultou no vetor K^{AG-FS} , e obtidas as estatísticas de média e desvio padrão.

Já na segunda linha da Tabela 6.10, a Matriz de avaliação inicial (MA) foi submetida aos testes com o AG, sendo mantida a ordem de preferência fornecida pelo decisor para os sete critérios, ou seja, $c_7 > c_1 > c_6 > c_4 > c_5 > c_2 > c_3$. Sendo assim, o vetor resultante (K^{AG-ED}) é uma otimização dos valores, respeitando-se a mesma ordem fornecida pelo decisor. E para o caso do vetor K^{AG} , não foi imposta a condição de respeitar-se nenhuma das duas ordens descritas anteriormente, ou seja, a opção de “Manter uma sequência pré-definida” ficou assinalada como “Não”.

Similarmente, a penúltima linha da Tabela 6.10 indica os testes feitos com a Matriz de Avaliação modificada ($+ c_{m+1}$), mantendo a ordem de preferência definida pelo decisor na análise com oito critérios ($c_7 > c_1 > c_6 > c_8 > c_4 > c_5 > c_2 > c_3$), o que resultou no vetor K^{AG-ED} . E, a última linha da Tabela 6.10, são os testes realizados com a mesma matriz, porém sem considerar a ordem de preferência, o que resultou no vetor K^{AG} .

Percebe-se que, de modo geral, as repetições realizadas com o AG nas combinações, de matrizes de avaliação com vetores de constante de escala, não tiveram grandes variações nos valores obtidos, tanto para os valores da constante de escala para cada vetor, quanto para a solução, que engloba as informações de “Custo para realização” e “Valor do Portfólio”. Isso deve-se, principalmente, às características do banco de dados, que tem alguns dos projetos/instâncias com alto custo para realização e que, conseqüentemente, dificilmente serão selecionados para compor o portfólio. Sendo assim, o algoritmo seleciona, quase sempre, os mesmos projetos, tendo assim pouca ou nenhuma variação no quesito “Custo para realização” e “Valor do Portfólio”.

Tabela 6.10 - Melhores valores encontrados e estatísticas básicas, em 10 repetições feitas do AG com os vetores de constante de escala combinados com as matrizes de avaliação

Matriz de Avaliação	Vetor de constante de escala (peso)	Estatísticas básicas	Valores de constante de escala (pesos) em cada critério								Solução	
			c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	Custo para realização	Valor Portfólio (V)
Matriz de Avaliação Inicial MA	K^{AG-FS}	Melhor vetor	0,012	0,001	0,017	0,02	0,016	0,475	0,457	---	878468	107,701276
		Média	0,1154	0,0037	0,1197	0,1203	0,1159	0,2629	0,259	---	878468	107,309324
		Desvio Padrão	0,0375	0,0062	0,0373	0,0366	0,0364	0,0771	0,0722	---	0	0,01491
	K^{AG-ED}	Melhor vetor	0,489	0	0	0	0	0,001	0,499	---	884361	117,889081
		Média	0,493	0	0	0,0005	0	0,0062	0,4976	---	884361	117,776978
		Desvio Padrão	0,002	0	0	0,0008	0	0,0043	0,0027	---	0	0,0691
	K^{AG}	Melhor vetor	0,89	0	0,055	0	0	0,009	0,044	---	884361	122,415692
		Média	0,8069	0,0033	0,0242	0,003	0,0053	0,0026	0,1527	---	884361	121,357653
		Desvio Padrão	0,0578	0,0073	0,0307	0,0069	0,0076	0,0047	0,0663	---	0	0,7676
Matriz de Avaliação modificada ($-a_i y_i = 0$)	K^{AG-FS}	Melhor vetor	0,001	0	0,014	0,021	0,001	0,851	0,107	---	813879	110,863443
		Média	0,0085	0	0,0291	0,0382	0,0169	0,7182	0,1139	---	813879	109,168051
		Desvio Padrão	0,0095	0	0,0118	0,0148	0,0099	0,2281	0,0191	---	0	1,1180
	K^{AG-ED}	Melhor vetor	0,332	0	0	0	0	0,332	0,334	---	813879	103,410895
		Média	0,3323	0	0	0,0004	0	0,3307	0,3342	---	813879	103,393995
		Desvio Padrão	0,0007	0	0	0,0013	0	0,0032	0,0004	---	0	0,0412
	K^{AG}	Melhor vetor	0,024	0	0,01	0	0	0,917	0,048	---	813879	112,142323
		Média	0,0102	0,0101	0,004	0,0035	0,0046	0,8724	0,0936	---	813879	111,334557
		Desvio Padrão	0,0135	0,0176	0,0050	0,0042	0,0078	0,0472	0,0512	---	0	0,6763
Matriz de Avaliação modificada ($+c_{m+1}$)	K^{AG-ED}	Melhor vetor	0,499	0	0	0	0	0	0,5	0	884361	117,90603
		Média	0,4917	0	0	0,0002	0,0001	0,0303	0,4951	0,0004	884361	117,682332
		Desvio Padrão	0,0053	0	0	0,0006	0,0003	0,0673	0,0049	0,001	0	0,1786
	K^{AG}	Melhor vetor	0,946	0	0,002	0	0	0	0,048	0	884361	123,350967
		Média	0,8915	0,0101	0,0113	0,0014	0	0,0608	0,0569	0	883516,6	122,289363
Desvio Padrão	0,0541	0,0121	0,0167	0,0027	0	0,1140	0,0451	0	1692,2983	0,8650		

Os resultados de todas as 10 repetições realizadas com o AG, em cada caso descrito na Tabela 6.10, estão expostos nos Anexos de L até S.

6.2.6. Caso II: Resultados

Um resumo geral das 13 soluções encontradas para o Caso II está apresentado na Tabela 6.11. Nesta tabela, é possível verificar as soluções para as combinações de cada Matriz de Avaliação com cada vetor de constante de escala (peso), obtidos no decorrer do processo de análise, seguindo-se as fases e procedimentos descritos no *framework* proposto.

Tabela 6.11 - Soluções encontradas para o Caso II

Matriz de Avaliação	Vetor de Constantes de escala	Ordem de preferência dos critérios	Solução		
			Projetos selecionados	Custo para realização	Valor do portfólio
MA	$K^{FS} = [0,012; 0,001; 0,017; 0,020; 0,016; 0,475; 0,457]$	$c_6 > c_7 > c_4 > c_3 > c_5 > c_1 > c_2$	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 24 25 26 27 28 29 30 31 32 33 34 35 36 38 39 40	878468	107,688880
	$K^{ED} = [0,260; 0,028; 0,014; 0,101; 0,057; 0,246; 0,289]$	$c_7 > c_1 > c_6 > c_4 > c_5 > c_2 > c_3$	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 24 25 26 27 28 29 30 31 32 33 34 35 36 38 39 40	878468	110,348306
	$K^{AG-FS} = [0,012; 0,001; 0,017; 0,02; 0,016; 0,475; 0,457]$	$c_6 > c_7 > c_4 > c_3 > c_5 > c_1 > c_2 *$	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 24 25 26 27 28 29 30 31 32 33 34 35 36 38 39 40	878468	107,701276
	$K^{AG-ED} = [0,489; 0,000; 0,000; 0,000; 0,001; 0,499]$	$c_7 > c_1 > c_6 > c_4 > c_5 > c_2 > c_3 *$	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40	884361	117,889081
	$K^{AG} = [0,890; 0,000; 0,055; 0,000; 0,000; 0,009; 0,044]$	$c_1 > c_3 > c_7 > c_6 \succ c_2 \succ c_4 \succ c_5 **$	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40	884361	122,415692

Continua...

Tabela 6.12 - Soluções encontradas para o Caso II (continuação)

MA modificada (- $a_i y_i = 0$)	$K^{FS} = [0,012;$ 0,001; 0,017; 0,020; 0,016; 0,475; 0,457]	$c_6 > c_7 > c_4 > c_3 > c_5 > c_1 > c_2$	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40	813879	105,146697
	$K^{ED} = [0,260;$ 0,028; 0,014; 0,101; 0,057; 0,246; 0,289]	$c_7 > c_1 > c_6 > c_4 > c_5 > c_2 > c_3$	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40	813879	97,911032
	$K^{AG-FS} =$ [0,001; 0,000; 0,014; 0,021; 0,001; 0,851; 0,107]	$c_6 > c_7 > c_4 > c_3 > c_5 > c_1 > c_2 *$	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40	813879	110,863443
	$K^{AG-ED} =$ [0,332; 0,000; 0,000; 0,000; 0,000; 0,332; 0,334]	$c_7 > c_1 > c_6 > c_4 > c_5 > c_2 > c_3 *$	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40	813879	103,410895
	$K^{AG} = [0,024;$ 0,000; 0,001; 0,000; 0,000; 0,917; 0,048]	$c_6 > c_7 > c_1 > c_3 \gtrsim c_2 \gtrsim c_4 \gtrsim c_5 **$	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40	813879	112,142323
MA modificada (+ c_{m+1})	$K^{ED} = [0,227;$ 0,025; 0,012; 0,088; 0,050; 0,215; 0,253; 0,126]	$c_7 > c_1 > c_6 > c_8 > c_4 > c_5 > c_2 > c_3$	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 24 25 26 27 28 29 30 31 32 33 34 35 36 38 39 40	878468	102,423679
	$K^{AG-ED} =$ [0,499; 0,000; 0,000; 0,000; 0,000; 0,000; 0,500; 0,000]	$c_7 > c_1 > c_6 > c_8 > c_4 > c_5 > c_2 > c_3 *$	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40	884361	117,90603
	$K^{AG} = [0,946;$ 0,000; 0,002; 0,000; 0,000; 0,000; 0,048; 0,000]	$c_1 > c_3 > c_7 \gtrsim c_2 \gtrsim c_4 \gtrsim c_5 \gtrsim c_6$ $\gtrsim c_8 **$	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40	884361	123,350967

* $c_j > c_{j+1}$ indica que o critério c_j é preferível ao critério c_{j+1} , desconsiderando o arredondamento dos valores;

** $c_j \gtrsim c_{j+1}$ indica que o critério c_j é preferível ou equivalente ao critério c_{j+1} .

Neste caso, novamente, ressalta-se que por conta do arredondamento dos valores de constante de escala (peso) de cada um dos critérios, às vezes não é possível verificar a ordem de preferência/importância dos mesmos, uma vez que se optou por representá-los com apenas quatro casas decimais. Porém, a ordem é sempre respeitada e mantida para os vetores que possuem essa condição, que no caso do K^{AG-FS} é respeitada a mesma ordem do K^{FS} e, similarmente, no caso do K^{AG-ED} é respeitada a mesma ordem do K^{ED} . E, também por conta do arredondamento dos valores, não foi possível identificar a ordem completa dos vetores K^{AG} .

Pelos dados descritos na Tabela 6.11, constata-se que os valores do portfólio obtidos a partir dos vetores otimizados pelo AG foram sempre superiores em relação ao seu respectivo vetor original. Por exemplo, os vetores K^{AG-ED} proporcionaram um “Valor do Portfólio” superior ao dos vetores K^{ED} , e o mesmo vale para os vetores K^{AG-FS} , que resultaram em um valor de portfólio maior do que o K^{FS} . Por sua vez, o vetor que não considera a ordem de importância/preferência pré-estabelecida, o K^{AG} , teve um resultado superior aos demais quanto ao “Valor do Portfólio”.

Embora os vetores obtidos pelo procedimento de AG tenham resultado em melhores valores de portfólio, o propósito maior de utilizar-se deste procedimento é proporcionar condições de serem calculados portfólios ótimos em situações em que o decisor não esteja apto, ou não deseja, efetuar a elicitación de magnitude dos pesos.

Por exemplo, considerando qualquer caso real, em que o decisor é uma pessoa experiente no assunto, tenha sua estrutura de preferências bem definida, e esteja disponível para que seja realizada a elicitación das constantes de escala, a solução obtida com este vetor (no caso K^{ED}) será altamente satisfatória. Porém, para um decisor que não tenha todas essas qualificações, é muito mais simples ele informar somente a ordem dos critérios e, a partir da otimização com o AG, obter uma solução com o vetor K^{AG-ED} .

Da mesma forma, o uso de registros históricos é extremamente valioso quando se trata de decisões recorrentes, como é o caso da decisão de concessão de crédito bancário. Estas informações podem ser utilizadas para classificar as instâncias em “sucesso” ou “insucesso”, com base em um algoritmo que pode atingir até 85% de acurácia, e até mesmo para informar qual é a probabilidade de “sucesso” de um determinado projeto, com base no treinamento das redes bayesianas. Uma outra informação relevante obtida com os registros históricos é o *ranking* dos atributos mais importantes que, por sua vez, pode vir a substituir a ordem de preferência fornecida pelo decisor, quando este não está disponível, ou então ser usado como uma informação útil para auxiliar o decisor no processo de ordenação e definição de pesos dos critérios.

7. CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

Esta pesquisa elaborou um modelo de análise de decisão multicritério para problemas de seleção de portfólio, empregando uma visão multidisciplinar, que englobou a aplicação de técnicas, métodos e ferramentas não somente da área de decisão multicritério, mas também de outras áreas do conhecimento, que estão intimamente ligadas com a área de PO, tais como: estatística; inteligência artificial; aprendizagem de máquina; computação evolucionária e otimização combinatória. Vale destacar que o levantamento bibliográfico realizado demonstrou a necessidade de mais estudos que relacionem a área de decisão multicritério com demais campos do conhecimento da PO, principalmente se tratando do problema de seleção de portfólio.

O resultado da pesquisa foi a criação de um *framework* de análise, constituído de diversas fases e procedimentos, que considera tanto a diversidade das possibilidades de casos, quanto a flexibilidade para a aplicação de diferentes métodos e/ou técnicas, porém com uma estrutura sólida e bem organizada dos possíveis caminhos a serem seguidos, rumo à melhores soluções para o problema de seleção de portfólio.

As justificativas que nortearam os objetivos desta pesquisa devem-se principalmente a dois fatores recorrentes nos dias atuais, no que tange o contexto da decisão: na dificuldade em estabelecer os parâmetros necessários para aplicação dos métodos, e no uso das experiências passadas como forma de aprendizado para tomada de melhores decisões no presente.

O *framework* desenvolvido fornece suporte a decisões que se enquadram em um, ou ambos, dos fatores acima mencionados, por meio da aplicação e implementação de técnicas que auxiliaram na determinação dos parâmetros, em específico as constantes de escala (pesos), e na predição da classificação de novos projetos, com base em registros históricos, resultando assim em soluções mais confiáveis, e até mesmo otimizadas.

Todas as quatro fases principais do *framework* que constituem o modelo proposto, bem como seus respectivos procedimentos, foram definidos com relação aos objetivos estabelecidos e também com relação aos possíveis métodos e/ou técnicas que podem ser utilizados. Como por exemplo, na Fase de Predição e no procedimento de Aprendizagem Supervisionada foram elencados vários algoritmos classificadores que podem ser utilizados, porém não necessariamente é obrigatório o uso de todos eles, tendo o analista/especialista a total liberdade de escolher o algoritmo que deseja utilizar.

Na Fase de Seleção e Otimização foi necessária a implementação dos algoritmos propostos em um *software*, desenvolvido em parceria com pesquisadores da Universidade

Federal de Pernambuco, cuja implementação e integração do algoritmo genético com o método de seleção de portfólio foi realizado pela autora. Foram realizados testes com o *software* desenvolvido, tanto com um problema hipotético, quanto em estudos de caso reais, obtendo-se assim resultados satisfatórios.

Vale ressaltar algumas considerações referentes ao modelo do *framework* proposto, bem como da aplicabilidade do mesmo, como seguem:

- Das premissas e limitações do modelo: os problemas a serem resolvidos pelo modelo são exclusivamente da problemática de seleção de portfólio, com uma única restrição;
- Da implementação do *software*: somente a fase de Seleção e Otimização foi implementada de forma integrada em um *software*;
- Dos estudos de caso: ambos os estudos de caso foram adaptados, com a introdução de uma única restrição hipotética, a fim de utilizar-se do modelo.

Por fim, embora visto que o modelo proposto no *framework* de análise encontra-se bem estruturado, com as etapas bem definidas, buscando abranger o maior número de possibilidades de casos para a aplicação em problemas reais de seleção de portfólio, muitos esforços ainda são necessários para aprimorar as contribuições que o modelo pode e poderá oferecer, principalmente no que diz respeito a implementação do modelo em um *software* que tenha um sistema integrado completo de todas as fases. Além desta necessidade, outras sugestões de pesquisas, que têm grande potencial para um futuro promissor do modelo, são:

- Possibilidade da introdução de mais de uma restrição no *software*, que implementa o problema de programação linear inteira considerado no modelo;
- Possibilidade de escolha de qual ou quais métodos ou métricas deseja-se efetuar a análise, de forma automática em um único *software*;
- Possibilidade de trabalhar com informações parciais e/ou incompletas, com relação ao desempenho de cada alternativa em cada critério;
- A aplicação do modelo proposto em outros estudos de casos, principalmente em situações reais que tenham os dados de registros históricos como, por exemplo, em instituições financeiras de fomento, instituições de ensino superior ou outras instituições que necessitem da aplicação de ferramentas para a seleção de portfólio.

REFERÊNCIAS

- ABBASIANJAHROMI, H.; RAJAIE, H. Developing a project portfolio selection model for contractor firms considering the risk factor. **Journal of Civil Engineering and Management**, v. 18, n. 6, p. 879-889, 2012.
- ABDOLLAHI, M.; ARVAN, M.; RAZMI, J. An integrated approach for supplier portfolio selection: Lean or agile? **Expert Systems with Applications**, v. 42, n. 1, p. 679-690, Jan 2015.
- ABEPRO. **Áreas e subáreas da Engenharia de Produção**. Disponível em: <<http://www.abepro.org.br/interna.asp?p=399&s=1&c=362>>.
- ABIYEV, R. H.; MENEKAY, M. Fuzzy portfolio selection using genetic algorithm. **Soft Computing**, v. 11, n. 12, p. 1157-1163, Oct 2007.
- AHMADI, S. H.; ALIZADEH SHANI, S. R. Ranking pharmaceuticals industry using SD-heuristics approach. **International Journal of Management and Business Research**, v. 6, n. 1, p. 23-34, 2016. 2016.
- ALI, R. et al. Fuzzy Evaluation Method for Project Profitability. In: SHAKHOVSKA, N. (Ed.). **Advances in Intelligent Systems and Computing**, Csit 2016, v.512, p.17-27, 2017.
- ALMEIDA, A. T. **Processo de decisão nas organizações: construindo modelos de decisão multicritério**. São Paulo: Atlas, 2013.
- ALMEIDA, A. T.; DUARTE, M. D. O. A multi-criteria decision model for selecting Project Portfolio with consideration being given to a new concept for synergies. **Pesquisa Operacional**, v. 31, n. 2, p. 301-318, 2011. 2011.
- ALMEIDA, A. T.; VETSCHERA, R.; ALMEIDA, J. A. Scaling issues in additive multicriteria portfolio analysis. In: **Decision Support Systems III-Impact of Decision Support Systems for Global Environments**. Springer, Cham, p. 131-140. 2014.
- ALMEIDA, J. A.; ALMEIDA, A. T.; COSTA, A. P. C. S. Portfolio selection of information systems projects using PROMETHEE V with c-optimal concept. **Pesquisa Operacional**, v. 34, n. 2, p. 275-299, 2014. 2014.
- ALPAYDIN, E. **Introduction to Machine Learning**. 2010.
- AN, D. et al. Mitigating pollution of hazardous materials from WEEE of China: Portfolio selection for a sustainable future based on multi-criteria decision making. **Resources, Conservation and Recycling**, v. 105, p. 198-210, 2015.
- AN, D. et al. Mitigating pollution of hazardous materials from WEEE of China: Portfolio selection for a sustainable future based on multi-criteria decision making. **Resources, Conservation and Recycling**, v. 105, p. 198-210, 2015.
- ANAGNOSTOPOULOS, K. P.; MAMANIS, G. A portfolio optimization model with three objectives and discrete variables. **Computers & Operations Research**, v. 37, n. 7, p. 1285-1297, Jul 2010.
- ANAGNOSTOPOULOS, K. P.; MAMANIS, G. The mean-variance cardinality constrained portfolio optimization problem: An experimental evaluation of five multiobjective evolutionary algorithms. **Expert Systems with Applications**, v. 38, n. 11, p. 14208-14217, Oct 2011.

ANDREICA, M. E. et al. A New Portfolio Selection Method Based on Interval Data. **Studies in Informatics and Control**, v. 19, n. 3, p. 253-262, Sep 2010.

ANDRIOSPOULOS, K.; NOMIKOS, N. Performance replication of the Spot Energy Index with optimal equity portfolio selection: Evidence from the UK, US and Brazilian markets. **European Journal of Operational Research**, v. 234, n. 2, p. 571-582, Apr 2014.

ARASTEH, A.; ALIAHMADI, A.; OMRAN, M. M. A Multi-stage Multi Criteria Model for Portfolio Management. **Arabian Journal for Science and Engineering**, v. 39, n. 5, p. 4269-4283, May 2014.

BACANIN, N.; TUBA, M.; IEEE. **Fireworks Algorithm Applied to Constrained Portfolio Optimization Problem**. p. 1242-1249, 2015.

BACANIN, N.; TUBA, M.; PELEVIC, B. Constrained portfolio selection using Artificial Bee Colony (ABC) algorithm. **Optimization**, v. 17, p. 18, 2014 2014.

BAHLOUL, S.; ABID, F. A combined analytic hierarchy process and goal programming approach to international portfolio selection in the presence of investment barriers. **International Journal of Multicriteria Decision Making**, v. 3, n. 1, p. 1-20, 2013. 2013.

BANCO CENTRAL DO BRASIL. **Recomendações de Basileia**, 2018. Disponível em: <<https://www.bcb.gov.br/fis/supervisao/basileia.asp>>. Acesso em: maio, 2018.

BANCO CENTRAL DO BRASIL. **Resolução nº 1.559, de 22/12/1988**. Disponível em: <<https://www.bcb.gov.br/pre/normativos/busca/normativo.asp?tipo=res&ano=1988&numero=1559>>. Acesso em: maio, 2018.

BANCO CENTRAL DO BRASIL. **Resolução nº 4.379, de 30/10/2014**. Disponível em: <<https://www.bcb.gov.br/pre/normativos/busca/normativo.asp?tipo=Resolu%C3%A7%C3%A3o&data=2014&numero=4379>>. Acesso em: maio, 2018.

BARAK, S.; ABESSI, M.; MODARRES, M. Fuzzy turnover rate chance constraints portfolio model. **European Journal of Operational Research**, v. 228, n. 1, p. 141-147, 2013.

BARATI, M. A.; MOHAMMADI, M.; NADERI, B. Multi-period fuzzy mean-semi variance portfolio selection problem with transaction cost and minimum transaction lots using genetic algorithm. **International Journal of Industrial Engineering Computations**, v. 7, n. 2, p. 217-228, 2016.

BELAID, A.; RAZMAK, J.; IEEE. Multi-Criteria Decision Support Systems: A Glorious History and a Promising Future. In: **2013 5th International Conference on Modeling, Simulation and Applied Optimization**, 2013.

BELTON, V.; STEWART, T. **Multiple criteria decision analysis: an integrated approach**. Springer Science & Business Media, 2002.

BHATTACHARYYA, R.; HOSSAIN, S. A.; KAR, S. Fuzzy cross-entropy, mean, variance, skewness models for portfolio selection. **Journal of King Saud University-Computer and Information Sciences**, v. 26, n. 1, p. 79-87, 2014.

BHATTACHARYYA, R.; KAR, S.; MAJUMDER, D. D. Fuzzy mean-variance-skewness portfolio selection models by interval analysis. **Computers & Mathematics with Applications**, v. 61, n. 1, p. 126-137, Jan 2011.

BHATTACHARYYA, R.; KUMAR, P.; KAR, S. Fuzzy R&D portfolio selection of interdependent projects. **Computers & Mathematics with Applications**, v. 62, n. 10, p. 3857-3870, Nov 2011.

BILBAO-TEROL, A. et al. Using TOPSIS for assessing the sustainability of government bond funds. **Omega-International Journal of Management Science**, v. 49, p. 1-17, Dec 2014.

BILBAO-TEROL, A.; ARENAS-PARRA, M.; CANAL-FERNANDEZ, V. Selection of Socially Responsible Portfolios using Goal Programming and fuzzy technology. **Information Sciences**, v. 189, p. 110-125, Apr 2012.

BILSEL, R. U. Portfolio selection under uncertainty: A multi - Criteria approach. In: **IIE Annual Conference. Proceedings**. Institute of Industrial and Systems Engineers (IISE), 2007.

BOGINSKI, V. et al. A network-based data mining approach to portfolio selection via weighted clique relaxations. **Annals of Operations Research**, v. 216, n. 1, p. 23-34, May 2014.

BORTOLUZZI, M. B. O; LÓPEZ, H. M. L; ALMEIDA, J. A; ALMEIDA, A. T. Modelo de decisão multicritério de seleção de portfólio para alocação de recursos em uma empresa do setor elétrico. In: **Simpósio Brasileiro de Pesquisa Operacional**, 49. Blumenau, 2017.

CADENAS, J. M., CARRILLO, J. V., GARRIDO, M. C., IVORRA, C., LIERN, V. Exact and heuristic procedures for solving the fuzzy portfolio selection problem. **Fuzzy optimization and Decision making**, v. 11, n. 1, p. 29-46 2012.

CADENAS, J. M. et al. GASRP-FP: A hybrid metaheuristic to solve fuzzy portfolio selection with non-financial goals. In: ALONSO, J. M.;BUSTINCE, H., *et al* (Ed.). **Proceedings of the 2015 Conference of the International Fuzzy Systems Association and the European Society for Fuzzy Logic and Technology**, v.89, p.784-791, 2015.

CALVO, C.; IVORRA, C.; LIERN, V. Finding socially responsible portfolios close to conventional ones. **International Review of Financial Analysis**, v. 40, p. 52-63 2014.

CHANG, T. J., MEADE, N., BEASLEY, J. E., SHARAIHA, Y. M. Heuristics for cardinality constrained portfolio optimisation. **Computers & Operations Research**, v. 27, n. 13, p. 1271-1302 2000.

CHEN, C. H.; HSIEH, C. Y; LEE, Y. C. Actionable Stock Portfolio Mining by Using Genetic Algorithms. **Journal of Information Science and Engineering**, v. 32, n. 6, p. 1657-1678, Nov 2013.

CHIEN, C. F.; HUYNH, N. T. An Integrated Approach for IC Design R&D Portfolio Decision and Project Scheduling and a Case Study. **Ieee Transactions on Semiconductor Manufacturing**, v. 31, n. 1, p. 76-86, Feb 2018.

CHO, K. K.; MOON, B. K. A method for selecting the optimal portfolio of performance improvement projects in a manufacturing system. **International Journal of Industrial Engineering-Theory Applications and Practice**, v. 13, n. 1, p. 61-70, Mar 2006.

CLEMEN, R. T.; SMITH, J. E. On the choice of baselines in multiattribute portfolio analysis: a cautionary note. **Decision Analysis**, v. 6, n. 4, p. 256-262, 2009.

COLAPINTO, C.; JAYARAMAN, R.; MARSIGLIO, S. Multi-criteria decision analysis with goal programming in engineering, management and social sciences: a state-of-the art review. **Annals of Operations Research**, v. 251, n. 1-2, p. 7-40, Apr 2017.

CONVERTINO, M.; CONVERTINO, N. Sequential portfolio decision model for epilepsy death risk reduction. In: 7th **International Congress on Environmental Modelling and Software**, iEMsS 2014. International Environmental Modelling and Software Society, 2014.

CORAZZA, M.; FASANO, G.; GUSSO, R. Portfolio selection with an alternative measure of risk: Computational performances of Particle Swarm Optimization and Genetic Algorithms. In: **Mathematical and statistical methods for actuarial sciences and finance**. Springer, Milano, 2012.

CRUZ, L. et al. Many-Objective Portfolio Optimization of Interdependent Projects with 'a priori' Incorporation of Decision-Maker Preferences. **Applied Mathematics & Information Sciences**, v. 8, n. 4, p. 1517-1531, Jul 2014.

CRUZ, L. et al. Multicriteria optimization of interdependent project portfolios with 'a priori' incorporation of decision maker preferences. In: LOPEZ, J. C. L.; ANDRADE, R. A. E., *et al* (Ed.). **Proceedings of the Fourth International Workshop on Knowledge Discovery, Knowledge Management and Decision Support**, v.51, 2013.

CRUZ-REYES, L. et al. An interactive decision support system framework for social project portfolio selection. In: **Recent Advances on Hybrid Intelligent Systems**. Springer, Berlin, Heidelberg, 2013. p. 377-391.

CRUZ-REYES, L. et al. Reduction of Decision Rules for Project Explanation on Public Project Portfolio. **International Journal of Combinatorial Optimization Problems and Informatics**, v. 6, n. 3, p. 5-21, Sep-Dec 2015.

CVÖRNJEK, N., BREZOCNIK, M., JAGRIC, T., PAPA, G. Comparison between single and multi objective genetic algorithm approach for optimal stock portfolio selection. 2014.

DAS, P.; BANERJEE, A. Meta optimization and its application to portfolio selection. In: **Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining**. ACM, 2011.

DASTKHAN, H.; GHARNEH, N. S.; GOLMAKANI, H. A linguistic-based portfolio selection model using weighted max-min operator and hybrid genetic algorithm. **Expert Systems with Applications**, v. 38, n. 9, p. 11735-11743, Sep 2011.

DEAN, J. **Big Data, Data Mining, and Machine Learning: Value Creation for Business Leaders and Practitioners**. John Wiley & Sons, 2014.

DEBNATH, A. et al. A Hybrid MCDM Approach for Strategic Project Portfolio Selection of Agro By-Products. **Sustainability**, v. 9, n. 8, Aug 2017.

DEHOUCHE, N. Non-profit project portfolio evaluation and selection: A multi-criteria approach. **International Journal of Applied Management Science**, v. 7, n. 4, p. 338-363, 2015..

DENG, G. F.; CHEN, C. L. Applying particle swarm optimization to solve portfolio selection problems. 2007.

DENG, G. F.; LIN, W. T.; LO, C. C. Markowitz-based portfolio selection with cardinality constraints using improved particle swarm optimization. **Expert Systems with Applications**, v. 39, n. 4, p. 4558-4566, Mar 2012.

DENG, X.; LI, R. J. Gradually tolerant constraint method for fuzzy portfolio based on possibility theory. **Information Sciences**, v. 259, p. 16-24, Feb 2014.

DETEMPLE, J. Portfolio Selection: A Review. **Journal of Optimization Theory and Applications**, v. 161, n. 1, p. 1-21, 2014.

DOBROVOLSKIENE, N.; TAMOSIUNIENE, R. An Index to Measure Sustainability of a Business Project in the Construction Industry: Lithuanian Case. **Sustainability**, v. 8, n. 1, Jan 2016a.

DOBROVOLSKIENE, N.; TAMOSIUNIENE, R. Sustainability-Oriented Financial Resource Allocation in a Project Portfolio through Multi-Criteria Decision-Making. **Sustainability**, v. 8, n. 5, May 2016b.

DOERNER, K. et al. Pareto ant colony optimization: A metaheuristic approach to multiobjective portfolio selection. **Annals of Operations Research**, v. 131, n. 1-4, p. 79-99, Oct 2004.

DOMBI, J.; ZSIROS, A. Learning multicriteria classification models from examples: Decision rules in continuous space. **European Journal of Operational Research**, v. 160, n. 3, p. 663-675, 2005.

DORFESHAN, Y. et al. Selecting project-critical path by a new interval type-2 fuzzy decision methodology based on MULTIMOORA, MOOSRA and TPOP methods. **Computers & Industrial Engineering**, v. 120, p. 160-178, Jun 2018.

DOUMPOS, M.; ZOPOUNIDIS, C. Computational intelligence techniques for multicriteria decision aiding: An overview. **Multicriteria Decision Aid and Artificial Intelligence**, p. 1-23, 2013.

DOUMPOS, M.; ZOPOUNIDIS, C. Financial modeling under multiple criteria. In: **Network Models in Economics and Finance**. Springer, Cham, 2014.

DRISSI, M.; OUMSIS, M.; ABOUTAJDINE, D. A Multi-Criteria Decision Framework for network selection over LTE and WLAN. **Engineering Applications of Artificial Intelligence**, v. 66, p. 113-127, Nov 2017.

DUTTA, S. et al. Fuzzy stochastic price scenario based portfolio selection and its application to BSE using genetic algorithm. **Applied Soft Computing**, v. 62, p. 867-891, Jan 2018.

EFTEKHARIAN, S. E.; SHOJAFAR, M.; SHAMSHIRBAND, S. 2-Phase NSGA II: An Optimized Reward and Risk Measurements Algorithm in Portfolio Optimization. **Algorithms**, v. 10, n. 4, Dec 2017.

EMMERICH, M. et al. Maximizing Consensus in Portfolio Selection in Multicriteria Group Decision Making. In: VARAJAO, J. E. Q.; CRUZCUNHA, M. M., *et al* (Ed.). **International Conference on Enterprise Information Systems/International Conference on Project Management/International Conference on Health and Social Care Information Systems and Technologies, Centeris/Projman / Hcist 2016**, v.100, p.848-855, 2016.

ENGAU, A. Tradeoff-based decomposition and decision-making in multiobjective programming. **European Journal of Operational Research**, v. 199, n. 3, p. 883-891, Dec 2009.

FARZI, S.; SHAVAZI, A. R.; PANDARI, A. R. Using Quantum-Behaved Particle Swarm Optimization for portfolio selection problem. **Int. Arab J. Inf. Technol.**, v. 10, n. 2, p. 111-119, 2013.

FENG, Y.; WANG, L.; LIU, X. H. Reliable Portfolio Selection Problem in Fuzzy Environment: An $m(\lambda)$ Measure Based Approach. **Algorithms**, v. 10, n. 2, Jun 2017.

FERNANDEZ, E. et al. Application of the non-outranked sorting genetic algorithm to public project portfolio selection. **Information Sciences**, v. 228, p. 131-149, Apr 2013.

FERNANDEZ, E. et al. Design and Solution of a Surrogate Model for Portfolio Optimization Based on Project Ranking. **Scientific Programming**, 2017.

FERNANDEZ, E. et al. Hybrid metaheuristic approach for handling many objectives and decisions on partial support in project portfolio optimisation. **Information Sciences**, v. 315, p. 102-122, Sep 2015.

- FOROOZESH, N.; TAVAKKOLI-MOGHADDAM, R.; MOUSAVI, S. M. Sustainable-supplier selection for manufacturing services: a failure mode and effects analysis model based on interval-valued fuzzy group decision-making. **International Journal of Advanced Manufacturing Technology**, v. 95, n. 9-12, p. 3609-3629, Apr 2018.
- FOULADGAR, M. M. et al. PROJECT PORTFOLIO SELECTION USING FUZZY AHP AND VIKOR TECHNIQUES. **Transformations in Business & Economics**, v. 11, n. 1, p. 213-231, 2012.
- FRANK, A. G. et al. A framework for decision-making in investment alternatives selection. **International Journal of Production Research**, v. 51, n. 19, p. 5866-5883, Oct 2013.
- FREITAS, A. A. **A survey of evolutionary algorithms for data mining and knowledge discovery**. In: Advances in evolutionary computing. Springer, Berlin, Heidelberg, 2003. p. 819-845.
- GABREL, V.; MURAT, C.; THIELE, A. Recent advances in robust optimization: An overview. **European Journal of Operational Research**, v. 235, n. 3, p. 471-483, Jun 2014.
- GAO, W., SHENG, H., WANG, J., WANG, S. Artificial Bee Colony Algorithm Based on Novel Mechanism for Fuzzy Portfolio Selection. **IEEE Transactions on Fuzzy Systems**, 2018.
- GIL, A. C. **Como elaborar projetos de pesquisa**. São Paulo, 2008.
- GILLI, M.; SCHUMANN, E. Heuristic optimisation in financial modelling. **Annals of operations research**, v. 193, n. 1, p. 129-158, 2012.
- GOLMAKANI, H. R.; FAZEL, M. Constrained Portfolio Selection using Particle Swarm Optimization. **Expert Systems with Applications**, v. 38, n. 7, p. 8327-8335, Jul 2011.
- GOLMOHAMMADI, A.; PAJOUTAN, M. Meta heuristics for dependent portfolio selection problem considering risk. **Expert Systems with Applications**, v. 38, n. 5, p. 5642-5649, May 2011.
- GOMEZ, M. A.; FLORES, C. X.; OSORIO, M. A. **Hybrid search for cardinality constrained portfolio optimization**. 2006.
- GRISHINA, N.; LUCAS, C. A.; DATE, P. Prospect theory-based portfolio optimization: an empirical study and analysis using intelligent algorithms. **Quantitative Finance**, v. 17, n. 3, p. 353-367, Mar 2017.
- GUO, S. et al. Fuzzy multi-period portfolio selection with different investment horizons. **European Journal of Operational Research**, v. 254, n. 3, p. 1026-1035, 2016.
- GUO, Y. et al. Balancing strategic contributions and financial returns: a project portfolio selection model under uncertainty. **Soft Computing**, p. 1-13, 2018.
- GUPTA, P. et al. Multi-criteria Portfolio Optimization Using Support Vector Machines and Genetic Algorithms. In: (Ed.). **Fuzzy Portfolio Optimization: Advances in Hybrid Multi-criteria Methodologies**. Berlin, Heidelberg: Springer Berlin Heidelberg, p.283-309, 2014. .
- GUPTA, P.; MEHLAWAT, M. K.; MITTAL, G. Asset portfolio optimization using support vector machines and real-coded genetic algorithm. **Journal of Global Optimization**, v. 53, n. 2, p. 297-315, Jun 2012.
- GUPTA, P.; MEHLAWAT, M. K.; SAXENA, A. A hybrid approach to asset allocation with simultaneous consideration of suitability and optimality. **Information Sciences**, v. 180, n. 11, p. 2264-2285, Jun 2010.

- HASSANLOU, K. A multi period portfolio selection using chance constrained programming. 2017.
- HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The elements of statistical learning**. 2.ed. New York: Springer series in statistics, 2008.
- HELMBOLD, D. P. et al. On-line portfolio selection using multiplicative updates. **Mathematical Finance**, v. 8, n. 4, p. 325-347, Oct 1998.
- HOCINE, A. et al. Optimizing renewable energy portfolios under uncertainty: A multi-segment fuzzy goal programming approach. **Renewable Energy**, v. 129, p. 540-552, 2018.
- HUANG, K. Y.; JANE, C. J. An automatic stock market forecasting and portfolio selection mechanism based on VPRS, ARX and grey system. In: **Asia-Pacific Services Computing Conference**, 2008. APSCC'08. IEEE. IEEE, 2008.
- HUANG, S. H.; LAI, S. H.; TAI, S. H. A Learning-Based Contrarian Trading Strategy via a Dual-Classifer Model. **Acm Transactions on Intelligent Systems and Technology**, v. 2, n. 3, 2011.
- HUANG, X. Mean-semivariance models for fuzzy portfolio selection. **Journal of computational and applied mathematics**, v. 217, n. 1, p. 1-8, 2008.
- HUANG, X. Portfolio selection with a new definition of risk. **European Journal of Operational Research**, v. 186, n. 1, p. 351-357, 2008b.
- HUANG, X. X.; ZHAO, T. Y. Project selection and scheduling with uncertain net income and investment cost. **Applied Mathematics and Computation**, v. 247, p. 61-71, Nov 2014.
- HUANG, X. X.; ZHAO, T. Y.; KUDRATOVA, S. Uncertain mean-variance and mean-semivariance models for optimal project selection and scheduling. **Knowledge-Based Systems**, v. 93, p. 1-11, Feb 2016.
- INFORMS. **What is Operations Research?**. Disponível em: < <https://www.informs.org/About-INFORMS/What-is-Operations-Research>>. Acesso em: 23 nov. 2016.
- JAHAN, M. V.; AKBARZADEH-T, M. R. Extremal optimization vs. learning automata: Strategies for spin selection in portfolio selection problems. **Applied Soft Computing**, v. 12, n. 10, p. 3276-3284, Oct 2012.
- JAHAN, M. V.; AKBARZADEH-TOTONCHI, M. R. From local search to global conclusions: Migrating spin glass-based distributed portfolio selection. **IEEE Transactions on Evolutionary Computation**, v. 14, n. 4, p. 591-601, 2010.
- JIANG, Z.; ZHANG, H.; SUTHERLAND, J. W. Development of multi-criteria decision making model for remanufacturing technology portfolio selection. **Journal of Cleaner Production**, v. 19, n. 17-18, p. 1939-1945, 2011.
- JOSHI, D.; KUMAR, S. Intuitionistic fuzzy entropy and distance measure based TOPSIS method for multi-criteria decision making. **Egyptian Informatics Journal**, v. 15, n. 2, p. 97-104, 2014.
- KANG, P. Y. et al. **Applying Hueristic Algorithms to Portfolio Selection Problem**. p. 323-329, 2015.
- KAR, M. B. et al. A new bi-objective fuzzy portfolio selection model and its solution through evolutionary algorithms. **Soft Computing**, p. 1-15, 2018..

KHALIDJI, M. et al. Dynamically weighted continuous ant colony optimization for biobjective portfolio selection using value-at-risk. In: **Modelling & Simulation**, 2009. AMS'09. Third Asia International Conference on. IEEE, 2009.

KHALILI-DAMGHANI, K. et al. A hybrid fuzzy rule-based multi-criteria framework for sustainable project portfolio selection. **Information Sciences**, v. 220, p. 442-462, Jan 2013.

KIMIAGARI, A. M.; NIKKHOLGH, R.; GHARAHKOZLI, H. Genetic algorithms for fuzzy multi-objective approach to portfolio selection. In: **Fuzzy Information Processing Society (NAFIPS)**, 2010 Annual Meeting of the North American. IEEE, 2010.

KIRIS, S.; USTUN, O. Fuzzy MCDM approach of stocks evaluation and portfolio selection. In: **MEC EurOPT 2010**. Presented at the 24th Mini EURO Conference "Continuous Optimization and Information-Based Technologies in the Financial Sector", Izmir, Turkey. 2010. 2010.

KLEINMUNTZ, D.N. "Resource allocation decisions", In: Edwards, W., Miles, R.F. and von Winterfeldt, D. (Eds), **Advances in Decision Analysis**, Cambridge University Press, New York, NY, pp. 400-418, 2007.

KOLLER, D.; FRIEDMAN, N. **Probabilistic graphical models: principles and techniques**. MIT press, 2009.

KUO, R. J.; HONG, C. W. Integration of Genetic Algorithm and Particle Swarm Optimization for Investment Portfolio Optimization. **Applied Mathematics & Information Sciences**, v. 7, n. 6, p. 2397-2408, Nov 2013.

LAROSE, D. T.; LAROSE, C. D. **Data mining and predictive analytics**. John Wiley & Sons, 2015.

LI, X. et al. A hybrid intelligent algorithm for portfolio selection problem with fuzzy returns. **Journal of Computational and Applied Mathematics**, v. 233, n. 2, p. 264-278, Nov 2009.

LI, X.; QIN, Z. F.; KAR, S. Mean-variance-skewness model for portfolio selection with fuzzy returns. **European Journal of Operational Research**, v. 202, n. 1, p. 239-247, Apr 2010.

LIAGKOURAS, K.; METAXIOTIS, K. **A New Fitness Guided Crossover Operator and Its Application for Solving the Constrained Portfolio Selection Problem**. p.171-187, 2015.

LIAO, X.; FANG, C. Selection of supplier portfolio in the presence of operational risk and disruption risk. In: **Industrial Engineering and Engineering Management (IEEM)**, 2015 IEEE International Conference on. IEEE, 2016.

LIN, C. C.; LIU, Y. T. Genetic algorithms for portfolio selection problems with minimum transaction lots. **European Journal of Operational Research**, v. 185, n. 1, p. 393-404, Feb 2008.

LIN, C. T. **New product portfolio selection using fuzzy logic**. p. 114-118 2007.

LIU, Y. J.; ZHANG, W. G. A multi-period fuzzy portfolio optimization model with minimum transaction lots. **European Journal of Operational Research**, v. 242, n. 3, p. 933-941, May 2015.

LIU, Y. J.; ZHANG, W. G.; XU, W. J. Fuzzy multi-period portfolio selection optimization models using multiple criteria. **Automatica**, v. 48, n. 12, p. 3042-3053, Dec 2012.

LIU, Y. J.; ZHANG, W. G.; ZHAO, X. J. Fuzzy multi-period portfolio selection model with discounted transaction costs. **Soft Computing**, v. 22, n. 1, p. 177-193, Jan 2018.

LONG, N. C. et al. Clustering stock data for multi-objective portfolio optimization. **International Journal of Computational Intelligence and Applications**, v. 13, n. 02, p. 1450011, 2014.

LOPES, Y. G.; ALMEIDA, A. T. A multicriteria decision model for selecting a portfolio of oil and gas exploration projects. **Pesquisa Operacional**, v. 33, n. 3, p. 417-441, 2013.

LOPES, Y. G.; ALMEIDA, A. T. Assessment of synergies for selecting a project portfolio in the petroleum industry based on a multi-attribute utility function. **Journal of Petroleum Science and Engineering**, v. 126, p. 131-140, Feb 2015.

LÓPEZ, H. M. L.; ALMEIDA, A. T. Project portfolio selection in an electric utility company using PROMETHEE V. **Production**, v. 24, n. 3, p. 559-571, 2014.

LOUKERIS, N., BOUTALIS, Y., ARAMPATZIS, A., LIVANIS, S., MALTOUDOGLOU, L. Computational intelligence in optimal portfolio selection - The PI model. In: **Information, Intelligence, Systems and Applications (IISA)**, 2015 6th International Conference on. IEEE 2016.

LOUKERIS, N., BOUTALIS, Y., LIVANIS, S., ARAMPATZIS, A., MALTOUDOGLOU, L. Hybrid Jordan Elman nets in portfolio selection. In: **Information, Intelligence, Systems and Applications (IISA)**, 2015 6th International Conference on. IEEE, 2016.

LOUKERIS, N.; BEKIROU, S.; ELEFTHERIADIS, I. The Intelligent Portfolio Selection Optimization System, (IPSOS). 2016.

LOUKERIS, N.; ELEFTHERIADIS, I. Further Higher Moments in Portfolio Selection and A Priori Detection of Bankruptcy, Under Multi-layer Perceptron Neural Networks, Hybrid Neuro-genetic MLPs, and the Voted Perceptron. **International Journal of Finance & Economics**, v. 20, n. 4, p. 341-361, 2015.

LOUKERIS, N.; ELEFTHERIADIS, I.; LIVANIS, E. A novel approach on hybrid Support Vector Machines into optimal portfolio selection. In: **Signal Processing and Information Technology (ISSPIT)**, 2013 IEEE International Symposium on. IEEE, 2013.

LOUKERIS, N.; ELEFTHERIADIS, I.; LIVANIS, E.. Portfolio selection in hybrid Radial Basis Functions. In: **Information, Intelligence, Systems and Applications**, IISA 2014, The 5th International Conference on. IEEE, 2014.

LOUKERIS, N.; ELEFTHERIADIS, I.; LIVANIS, E.. The Portfolio Heuristic Optimisation System (PHOS). **Computational Economics**, v. 48, n. 4, p. 627-648, 2016.

MAGOC, T.; MODAVE, F. The optimality of non-additive approaches for portfolio selection. **Expert Systems with Applications**, v. 38, n. 10, p. 12967-12973, Sep 2011.

MARTÍNEZ-VEGA, D. A. et al. Modeling and project portfolio selection problem enriched with dynamic allocation of resources. In: **Fuzzy Logic Augmentation of Neural and Optimization Algorithms: Theoretical Aspects and Real Applications**. Springer, Cham 2018.

MARTINS, C. L., LÓPEZ, H. M. L., DE ALMEIDA, A. T., ALMEIDA, J. A., & BORTOLUZZI, M. B. D. O.. An MCDM project portfolio web-based DSS for sustainable strategic decision making in an electricity company. **Industrial Management & Data Systems**, 117(7), 1362-1375, 2017.

MASHAYEKHI, Z.; OMRANI, H. An integrated multi-objective Markowitz-DEA cross-efficiency model with fuzzy returns for portfolio selection problem. **Applied Soft Computing**, v. 38, p. 1-9, Jan 2016.

MASTORAKIS, K.; SISKOS, E. Value focused pharmaceutical strategy determination with multicriteria decision analysis techniques. **Omega-International Journal of Management Science**, v. 59, p. 84-96, Mar 2016.

MEHLAWAT, M. K. Behavioral optimization models for multicriteria portfolio selection. **Yugoslav Journal of Operations Research**, v. 23, n. 2, 2013.

METAWA, N.; HASSAN, M. K.; ELHOSENY, M. Genetic algorithm based model for optimizing bank lending decisions. **Expert Systems with Applications**, v. 80, p. 75-82, Sep 2017.

MISHRA, A. R.; RANI, P.; JAIN, D. INFORMATION MEASURES BASED TOPSIS METHOD FOR MULTICRITERIA DECISION MAKING PROBLEM IN INTUITIONISTIC FUZZY ENVIRONMENT. **Iranian Journal of Fuzzy Systems**, v. 14, n. 6, p. 41-63, Dec 2017.

MORADI, M.; SADOLLAH, A.; ESKANDAR, H. The application of water cycle algorithm to portfolio selection. **Economic Research-Ekonomska Istrazivanja**, v. 30, n. 1, p. 1277-1299, Sep 2017.

MORTON, A. Measurement issues in the evaluation of projects in a project portfolio. **European Journal of Operational Research**, v. 245, n. 3, p. 789-796, 2015.

MORTON, A.; KEISLER, J. M.; SALO, A. Multicriteria portfolio decision analysis for project selection. In: **Multiple Criteria Decision Analysis**. Springer, New York, NY, 2016. p. 1269-1298. 2016.

NAJAFI, A. A.; MUSHAKHIAN, S. Multi-stage stochastic mean-semivariance-CVaR portfolio optimization under transaction costs. **Applied Mathematics and Computation**, v. 256, p. 445-458, Apr 2015.

NEJAD, M. K. Clustering stock exchange data by using evolutionary algorithms for portfolio management. **European Research Studies**, v. 17, n. 4, p. 55, 2014.

OSEI-BRYSON, K.; BARCLAY, C. **Knowledge Discovery Process and Methods**. CRC Press: Taylor & Francis Group, 2015.

PAIVA, F. D. et al. Decision-Making for Financial Trading: A Fusion Approach of Machine Learning and Portfolio Selection. **Expert Systems with Applications**, 2018.

PATARI, E. et al. Comparison of the multicriteria decision-making methods for equity portfolio selection: The US evidence. **European Journal of Operational Research**, v. 265, n. 2, p. 655-672, Mar 2018.

PENG, Z.; LUO, W.; DENG, A. A hybrid fuzzy outranking approach for group decision-making under uncertainty. In: **Intelligent Transportation, Big Data & Smart City (ICITBS)**, 2018 International Conference on. IEEE, p. 596-600. 2018.

POUYA, A. R.; SOLIMANPUR, M.; REZAEI, M. J. Solving multi-objective portfolio optimization problem using invasive weed optimization. **Swarm and Evolutionary Computation**, v. 28, p. 42-57, Jun 2016.

PRELIPCEAN, G.; BOSCOIANU, M.; POPOVICIU, N. The role of predictability of financial series in emerging market applications. In: **Proceedings of the 9th WSEAS International Conference on Mathematics & Computers in Business and Economics (MCBE'80)**. p. 203-208. 2008.

- RABBANI, M.; BAJESTANI, M. A.; KHOSHKHOU, G. B. A multi-objective particle swarm optimization for project selection problem. **Expert Systems with Applications**, v. 37, n. 1, p. 315-321, Jan 2010.
- RADIUKYNIENE, I.; ZILINSKAS, A. Evolutionary methods for multi-objective portfolio optimization. In: AO, S. I.; GELMAN, L., *et al* (Ed.). **World Congress on Engineering 2008**, 2008.
- RAZI, F. F., ESHLAGHY, A. T., NAZEMI, J., ALBORZI, M., POOREBRAHIMI, A. A hybrid grey-based fuzzy C-means and multiple objective genetic algorithms for project portfolio selection. **International Journal of Industrial and Systems Engineering**, v. 21, n. 2, p. 154-179, 2015.
- RAZI, F. F., ESHLAGHY, A. T., NAZEMI, J., ALBORZI, M., POOREBRAHIMI, A.. A hybrid Promethee II based hierarchical clustering and NSGA-II for project portfolio selection. *Advances in Environmental Biology*, p. 1954-1962, 2014.
- RAZI, F. F. A hybrid grey relational analysis and nondominated sorting genetic algorithm-II for project portfolio selection. **Advances in Operations Research**, 2014.
- RAZI, F. F.; SHAHABI, V. Forming the stock optimized portfolio using model Grey based on C5 and the Shuffled frog leap algorithm. **Journal of Statistics & Management Systems**, v. 19, n. 3, p. 397-421, 2016.
- RAZI, F. F.; SHARIAT, S. H. A hybrid grey based artificial neural network and C&R tree for project portfolio selection. **Benchmarking-an International Journal**, v. 24, n. 3, p. 651-665, 2017.
- REBIASZ, B. *et al*. Hybrid framework for investment project portfolio selection. In: (Ed.). **Federated Conference on Computer Science and Information Systems, 2014**, v.2, 2014.
- REBIASZ, B.; GAWEL, B.; SKALNA, I. Fuzzy Multi-attribute Evaluation of Investments. In: GANZHA, M.; MACIASZEK, L., *et al* (Ed.). **2013 Federated Conference on Computer Science and Information Systems**, 2013.
- RELICH, M. A knowledge-based system for new product portfolio selection. In: **New Frontiers in Information and Production Systems Modelling and Analysis**. Springer, Cham, p. 169-187 2016.
- RONYASTRA, I. M.; GUNARTA, I. K.; CIPTOMULYONO, U. A multi criteria decision analysis for reinvestment action portfolio selection problem in an Indonesian real estate company. In: PUJAWAN, N.; CIPTOMULYONO, U., *et al* (Ed.). **Industrial Engineering and Service Science 2015, Iess 2015**, v.4, p.558-567. 2015.
- ROODPOSHTI, F. R. *et al*. Optimization of the mutual-fund portfolio of Tehran Stock Exchange using artificial neural networks and Genetic Algorithm. 2016.
- RUIZ-TORRUBIANO, R.; SUAREZ, A. Hybrid Approaches and Dimensionality Reduction for Portfolio Selection with Cardinality Constraints. **Ieee Computational Intelligence Magazine**, v. 5, n. 2, p. 92-107, May 2010.
- RUSSELL, B.; BINDER, J.; KOLLER, D.; KANAZAWA, K. Local learning in probabilistic networks with hidden variables. In: **Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence**, 1146–1152. San Francisco: Morgan Kaufmann, 1995.
- RUSSELL, S.; NORVIG, P. **Inteligência artificial**. Elsevier, 2004.
- SEYEDHOSSEINI, S. M.; ESFAHANI, M. J.; GHAFFARI, M. A novel hybrid algorithm based on a harmony search and artificial bee colony for solving a portfolio optimization problem using a mean-semi variance approach. **Journal of Central South University**, v. 23, n. 1, p. 181-188, Jan 2016.

SHAKHSI-NIAEI, M. et al. Application of genetic and differential evolution algorithms on selecting portfolios of projects with consideration of interactions and budgetary segmentation. **International Journal of Operational Research**, v. 22, n. 1, p. 106-128, 2015.

SHAKHSI-NIAEI, M.; TORABI, S. A.; IRANMANESH, S. H. A comprehensive framework for project selection problem under uncertainty and real-world constraints. **Computers & Industrial Engineering**, v. 61, n. 1, p. 226-237, 2011.

SHAYANFAR, E. et al. Prioritizing Interrelated Road Projects Using Metaheuristics. **Journal of Infrastructure Systems**, v. 22, n. 2, Jun 2016.

SIRISRISAKULCHAI, J. et al. Portfolio Optimization of Financial Returns Using Fuzzy Approach with NSGA-II Algorithm. **Journal of Advanced Computational Intelligence and Intelligent Informatics**, v. 19, n. 5, p. 619-623, Sep 2015.

SMITH-PERERA, A. et al. A Project Strategic Index proposal for portfolio selection in electrical company based on the Analytic Network Process. **Renewable & Sustainable Energy Reviews**, v. 14, n. 6, p. 1569-1579, Aug 2010.

SOLA, A. V. H.; MOTA, C. M. D. A multi-attribute decision model for portfolio selection aiming to replace technologies in industrial motor systems. **Energy Conversion and Management**, v. 57, p. 97-106, May 2012.

SOLIMANPUR, M.; MANSOURFAR, G.; GHAYOUR, F. Optimum portfolio selection using a hybrid genetic algorithm and analytic hierarchy process. **Studies in Economics and Finance**, v. 32, n. 3, p. 379-394, 2015.

SONG, Q.; LIU, A.; YANG, S. Y. Stock portfolio selection using learning-to-rank algorithms with news sentiment. **Neurocomputing**, v. 264, p. 20-28, 2017.

SOUREH, H. M.; AMANOLLAHI, G. F. Comparative evaluation of fuzzy logic and genetic algorithms models for portfolio optimization. **Management Science Letters**, v. 7, n. 5, p. 247-254, 2017.

SPANCHER, S. **Accelerating Discovery: mining unstructure information for hypothesis generation**. CRC Press: Taylor & Francis Group, 2016.

STEUER, R. E.; NA, P. Multiple criteria decision making combined with finance: A categorized bibliographic study. **European Journal of Operational Research**, v. 150, n. 3, p. 496-515, Nov 2003.

STEWART, T. J. A multi-criteria decision support system for r&d project selection. **Journal of the operational Research Society**, v. 42, n. 1, p. 17-26 1991.

STEWART, T. J.. Multiple objective project portfolio selection based on reference points. **Journal of Business Economics**, v. 86, n. 1-2, p. 23-33, 2016.

STRANG, K. D. Portfolio Selection Methodology for a Nuclear Project. **Project Management Journal**, v. 42, n. 2, p. 81-93, Mar 2011.

STRUMBERGER, I.; BACANIN, N.; TUBA, M. Constrained Portfolio Optimization by Hybridized Bat Algorithm. In: ALDABASS, D.;ACHALAKUL, T., *et al* (Ed.). **2016 7th International Conference on Intelligent Systems, Modelling and Simulation**, p.83-88, 2016.

STUMMER, C.; KIESLING, E. A MULTICRITERIA DECISION SUPPORT SYSTEM FOR COMPETENCE-DRIVEN PROJECT PORTFOLIO SELECTION. **International Journal of Information Technology & Decision Making**, v. 8, n. 2, p. 379-401, Jun 2009.

SUGIYAMA, M. **Introduction to statistical machine learning**. Morgan Kaufmann, 2015.

SUN, M. R.; WEI, C. An artificial bee colony algorithm for fuzzy portfolio model with concave transaction costs. In: **Management Science and Engineering (ICMSE)**, 2013 International Conference on. IEEE, p. 400-405. 2013.

TAKEDA, A.; GOTOH, J. Y.; SUGIAMA, M. Support vector regression as conditional value-at-risk minimization with application to financial time-series analysis. In: **Machine Learning for Signal Processing (MLSP)**, 2010 IEEE International Workshop on. IEEE, p. 118-123. 2010.

TALEBI, A.; MOLAEI, M. A.; SHEIKH, M. J. **Performance Investigation and Comparison of Two Evolutionary Algorithms in Portfolio Optimization: Genetic and Particle Swarm Optimization**. 430-437, 2010.

TANG, W. S.; WANG, Y. Q.; IEEE. Intelligent method for dynamic portfolio selection with probability criterion. In: (Ed.). **2004 Ieee International Conference on Systems, Man & Cybernetics, Vols 1-7**, p.3323-3327, 2004.

TAVANA, M.; KHALILI-DAMGHANI, K.; ABTAHI, A. R. A fuzzy multidimensional multiple-choice knapsack model for project portfolio selection using an evolutionary algorithm. **Annals of Operations Research**, v. 206, n. 1, p. 449-483, Jul 2013.

TCHANGANI, A. P.; PÉRÈS, F. Elicitation, evaluation and recommendation in multi-criteria decision making process. In: **Control, Decision and Information Technologies (CoDIT)**, 2017 4th International Conference on. IEEE, p. 1107-1112. 2017.

TOFIGHIAN, A. A. et al. Multi-period project portfolio selection under risk considerations and stochastic income. **Journal of Industrial Engineering International**, p. 1-14, 2018.

TOFIGHIAN, A. A.; NADERI, B. Modeling and solving the project selection and Scheduling. **Computers & Industrial Engineering**, v. 83, p. 30-38, May 2015.

TUBA, M.; BACANIN, N. Artificial Bee Colony Algorithm Hybridized with Firefly Algorithm for Cardinality Constrained Mean-Variance Portfolio Selection Problem. **Applied Mathematics & Information Sciences**, v. 8, n. 6, p. 2831-2844, Nov 2014.

VAN DYK, E.; SMITH, D. G. R&D portfolio selection by using qualitative pairwise comparisons. **Omega**, v. 18, n. 6, p. 583-594, 1990.

VARMA, K.; KUMAR, K. S. CRITERIA ANALYSIS AIDING PORTFOLIO SELECTION USING DEMATEL. In: RAJESH, R.;GANESH, K., *et al* (Ed.). **International Conference on Modelling Optimization and Computing**, v.38, 2012. p.3649-3661.

VERCHER, E.; BERMUDEZ, J. D. Portfolio optimization using a credibility mean-absolute semi-deviation model. **Expert Systems with Applications**, v. 42, n. 20, p. 7121-7131, Nov 2015.

VERNIC, R. Optimal investment with a constraint on ruin for a fuzzy discrete-time insurance risk model. **Fuzzy Optimization and Decision Making**, v. 15, n. 2, p. 195-217, Jun 2016.

VETSCHERA, R.; ALMEIDA, A. T. A PROMETHEE-based approach to portfolio selection problems. **Computers & Operations Research**, v. 39, n. 5, p. 1010-1020, 2012.

- VILKKUMAA, E. et al. Scenario-based portfolio model for building robust and proactive strategies. **European Journal of Operational Research**, v. 266, n. 1, p. 205-220, Apr 2018.
- WANG, B.; WANG, S. M.; WATADA, J. Fuzzy-Portfolio-Selection Models With Value-at-Risk. **Ieee Transactions on Fuzzy Systems**, v. 19, n. 4, p. 758-769, Aug 2011.
- WANG, Y. Q.; IEEE. **Intelligent method for solving optimal strategy of dynamic portfolio selection with credibility criterion**, 3325-3329. 2006.
- WATSON, C. E. et al. The application of portfolio selection to fuel channel inspection in advanced gas-cooled reactors. **Nuclear Engineering and Design**, v. 328, p. 145-153, Mar 2018.
- WITTEN, I. H.; FRANK, E. Weka. **Data Mining and Knowledge Discovery Handbook**, p. 1305-1314, 2005.
- WOODSIDE-ORIAKHI, M.; LUCAS, C.; BEASLEY, J. E. Heuristic algorithms for the cardinality constrained efficient frontier. **European Journal of Operational Research**, v. 213, n. 3, p. 538-550, Sep 2011.
- WU, Y. et al. An intuitionistic fuzzy multi-criteria framework for large-scale rooftop PV project portfolio selection: Case study in Zhejiang, China. **Energy**, v. 143, p. 295-309, 2018.
- XIA, Y. et al. A model for portfolio selection with order of expected returns. **Computers & Operations Research**, v. 27, n. 5, p. 409-422, 2000.
- XIA, Y.; WANG, S.; DENG, X. A compromise solution to mutual funds portfolio selection with transaction costs. **European Journal of Operational Research**, v. 134, n. 3, p. 564-581, 2001.
- XIDONAS, P. et al. Environmental corporate responsibility for investments evaluation: an alternative multi-objective programming model. **Annals of Operations Research**, v. 247, n. 2, p. 395-413, Dec 2016.
- XU, F.; CHEN, W. A portfolio model with quadratic subsection concave transaction costs based on PSO. In: **Control and Decision Conference**, 2009. CCDC'09. Chinese. IEEE, p. 1895-1898. 2009.
- YAAKOB, A. M.; GEGOV, A.; RAHMAN, S. F. A. Selection of alternatives using fuzzy networks with rule base aggregation. **Fuzzy Sets and Systems**, v. 341, p. 123-144, Jun 2018.
- YAAKOB, A. M.; SERGUIIEVA, A.; GEGOV, A. FN-TOPSIS: Fuzzy Networks for Ranking Traded Equities. **Ieee Transactions on Fuzzy Systems**, v. 25, n. 2, p. 315-332, Apr 2017.
- YAAKOB, S. B.; WATADA, J. A hybrid particle swarm optimization approach to mixed integer quadratic programming for portfolio selection problems. **International Journal of Simulation-Systems, Science & Technology**, v. 11, n. 5, 2010.
- YAN, W.; MIAO, R.; LI, S. R. Multi-period semi-variance portfolio selection: Model and numerical solution. **Applied Mathematics and Computation**, v. 194, n. 1, p. 128-134, Dec 2007.
- YANG, C. H.; LEE, K. C.; CHEN, H. C. Incorporating carbon footprint with activity-based costing constraints into sustainable public transport infrastructure project decisions. **Journal of Cleaner Production**, v. 133, p. 1154-1166, Oct 2016.
- YU, L.; WANG, S. Y.; LAI, K. K. Multi-Attribute Portfolio Selection with Genetic Optimization Algorithms. **Infor**, v. 47, n. 1, p. 23-30, Feb 2009.

YU, T. Y. et al. Generating effective defined-contribution pension plan using simulation optimization approach. **Expert Systems with Applications**, v. 39, n. 3, p. 2684-2689, Feb 2012.

YUEN, K. K. F. Towards a genetic direct least squares prioritization method for pairwise reciprocal matrices: Application of asset portfolio selection. In: **Machine Learning and Cybernetics (ICMLC)**, 2016 International Conference on. IEEE, p. 275-279. 2017.

ZAKI, Mohammed J.; MEIRA JR, Wagner; MEIRA, Wagner. **Data mining and analysis: fundamental concepts and algorithms**. Cambridge University Press, 2014.

ZHANG, Q.; HUANG, X. X.; TANG, L. M. Optimal multinational capital budgeting under uncertainty. **Computers & Mathematics with Applications**, v. 62, n. 12, p. 4557-4567, Dec 2011.

ZHANG, R. et al. Learning to Select Supplier Portfolios for Service Supply Chain. **Plos One**, v. 11, n. 5, May 2016.

ZHANG, W. G. et al. Evaluating methods of investment project and optimizing models of portfolio selection in fuzzy uncertainty. **Computers & Industrial Engineering**, v. 61, n. 3, p. 721-728, Oct 2011.

ZHENG, J.; CAILLOUX, O.; MOUSSEAU, V. Constrained Multicriteria Sorting Method Applied to Portfolio Selection. In: BRAFMAN, R. I.; ROBERTS, F. S., *et al* (Ed.). **Algorithmic Decision Theory**, v.6992, p.331-343. 2011.

ZOPOUNIDIS, C.; DOUMPOS, M. Multicriteria decision systems for financial problems. **Top**, v. 21, n. 2, p. 241-261, Jul 2013.

ZOPOUNIDIS, C.; DOUMPOS, M.; ZANAKIS, S. Stock evaluation using a preference disaggregation methodology. **Decision Sciences**, v. 30, n. 2, p. 313-336, Spr 1999.

ZULOAGA, M. S.; MOSER, B. R. Optimizing Resource Allocation in a Portfolio of Projects Related to Technology Infusion Using Heuristic and Meta-Heuristic Methods. In: KOCAOGLU, D. F.; ANDERSON, T. R., *et al* (Ed.). **2017 Portland International Conference on Management of Engineering and Technology**, 2017.

ANEXO A – Exemplo do arquivo de importação dos dados

	A	B	C	D	E		A	B	C	D
1	Criterio	X1	X2	X3		1	nome	cost		
2	Peso	0,3	0,5	0,2		2	limite	560		
3	tipo	1	1	1		3	op	<=		
4	a					4	a1	120		
5	b					5	a2	210		
6	c					6	a3	100		
7	Alternativa					7	a4	230		
8	A1	1	2	7		8	a5	150		
9	A2	3	4	4		9				
10	A3	2	3	2		10				
11	A4	4	7	9		11				
12	A5	6	9	4		12				
13						13				

ANEXO B – Exemplo do arquivo de saída do algoritmo genético

```

outputAG.txt - Bloco de notas
Arquivo Editar Formatar Exibir Ajuda
GENETIC ALGORITHM OUTPUT
-----

Run Information
-----
Chromosome Size: 3
Keep the gene sequence: Yes (1 0 2 )
Population Size: 5
Manually added chromosome:
Selection Function: Select Function
Genetic Operators:
  Simple Crossover - Rate: 0,8
  Binary Mutation - Rate: 0,1
  Crossover Points: 1
Termination Criteria:
  Number of generation: 10
Evaluation Function: Maximization

Initial Population (normalized): Chromosome and Fitness Value
  0,4190 0,4460 0,1330 - Fitness:19,6806089823356
  0,3720 0,5270 0,1000 - Fitness:20,3265330110627
  0,1890 0,7420 0,0680 - Fitness:22,2806046341651
  0,3850 0,4320 0,1820 - Fitness:19,9529939470226
  0,3260 0,4720 0,2010 - Fitness:20,576082749194

Initial Population (normalized and ordered): Chromosome and Fitness Value
  0,1890 0,7420 0,0680 - Fitness:22,2806046341651
  0,3260 0,4720 0,2010 - Fitness:20,576082749194
  0,3720 0,5270 0,1000 - Fitness:20,3265330110627
  0,3850 0,4320 0,1820 - Fitness:19,9529939470226
  0,4190 0,4460 0,1330 - Fitness:19,6806089823356

Output of each Iteration
-----
Selected Parents:
  0,3260 0,4720 0,2010
  0,3850 0,4320 0,1820

Created children: Chromosome and Fitness Value
  0,3460 0,4590 0,1930 - Fitness:20,358467293891
  0,3630 0,4460 0,1890 - Fitness:20,1826597872053

New Population (normalized and ordered): Chromosome and Fitness Value
  0,1890 0,7420 0,0680 - Fitness:22,2806046341651
  0,3260 0,4720 0,2010 - Fitness:20,576082749194
  0,3460 0,4590 0,1930 - Fitness:20,358467293891
  0,3720 0,5270 0,1000 - Fitness:20,3265330110627
  0,3630 0,4460 0,1890 - Fitness:20,1826597872053

[... ] continue for each iteration, until the last one [...]
-----
Selected Parents:
  0,2870 0,4990 0,2120
  0,2590 0,5190 0,2210

Created children: Chromosome and Fitness Value
  0,2790 0,5050 0,2150 - Fitness:21,0658854244194
  0,2670 0,5140 0,2180 - Fitness:21,1889346078421

New Population (normalized and ordered): Chromosome and Fitness Value
  0,1890 0,7420 0,0680 - Fitness:22,2806046341651
  0,2360 0,5350 0,2270 - Fitness:21,5074906994987
  0,2590 0,5190 0,2210 - Fitness:21,2734752585508
  0,2670 0,5140 0,2180 - Fitness:21,1889346078421
  0,2700 0,5120 0,2170 - Fitness:21,1575358188602

AG - version 1.1 (2018). Developed by students of PPGEPS(PUCPR) and CDSID(UFPE)

```

ANEXO C – Dados do problema do estudo de Caso I

Projetos	Impacto no Resultado	Alinhamento Estratégico	Melhoria de Indicadores Regulados	Contribuição para Satisfação dos Clientes	Probabilidade de Sucesso	Complexidade	Custo/Consumo
AM01	2	3	2	1	0,35	2	172018877,00
AM02	2	2	1	1	0,85	1	216278053,00
AM03	3	3	3	3	0,6	1	296494327,00
AM04	3	2	3	3	0,6	3	116691864,00
AM05	1	3	2	3	0,85	2	80254837,00
AM06	2	3	1	1	0,35	3	256396295,00
AM07	1	3	2	3	0,6	2	305070741,00
AM08	2	3	1	1	0,85	2	314741806,00
AM09	1	1	1	1	0,6	1	134817027,00
AM10	1	2	2	2	0,6	2	112186121,00
AM11	1	3	1	1	0,85	2	130684093,00
AM12	1	3	1	1	0,6	1	222684021,00
AM13	1	1	2	1	0,6	2	259829384,00
AM14	2	1	2	1	0,6	2	162279468,00
AM15	2	1	2	1	0,6	1	232890098,00
AM16	2	2	2	1	0,6	1	120372849,00
AM17	2	2	2	1	0,6	2	209612840,00
RR01	1	1	1	2	0,6	3	148564320,00
RR02	2	2	2	1	0,85	3	30692596,00
RR03	2	3	3	3	0,6	1	80892834,00
RR04	2	2	1	1	0,85	1	271544067,00
RR05	3	2	2	2	0,85	2	296802038,00
RR06	1	2	2	1	0,6	3	296665873,00
RR07	2	3	2	1	0,6	1	149901060,00
RR08	1	1	2	1	0,6	1	270487321,00
RR09	1	2	2	1	0,6	2	92613326,00
RR10	2	2	3	3	0,85	1	249432460,00
RR11	1	2	3	3	0,85	3	179749576,00
RR12	2	3	2	2	0,35	3	134412969,00
RR13	2	1	1	1	0,6	2	45339168,00
RR14	1	3	2	1	0,6	2	279959862,00
RR15	2	2	3	3	0,6	1	221479668,00
RR16	3	2	2	1	0,6	2	287823097,00
RR17	2	2	1	3	0,85	2	289154585,00
RR18	2	2	3	2	0,85	2	78256397,00
RR19	3	2	2	1	0,6	1	235046108,00
RR20	3	3	3	1	0,6	1	118754829,00
RR21	1	1	2	2	0,85	3	60287968,00
RR22	2	2	2	1	0,85	2	98767616,00
RR23	1	3	1	3	0,85	2	140931304,00
SC01	1	3	1	3	0,85	2	307422357,00
SC02	1	2	2	2	0,85	2	295998994,00
SC03	1	3	2	3	0,6	2	52853394,00
SC04	1	2	1	2	0,85	1	54771828,00
SC05	1	2	1	1	0,85	3	297145718,00
SC06	1	2	1	1	0,85	3	75912688,00
SC07	1	1	1	1	0,6	3	161234532,00
SC08	1	1	1	1	0,85	3	118260306,00
SC09	1	3	1	1	0,85	3	275922980,00
SC10	1	3	1	2	0,6	3	202665791,00
SC11	1	3	1	1	0,85	3	55230378,00
SC12	1	3	1	3	0,85	2	98009589,00
SC13	2	3	1	1	0,85	3	187218184,00
SC14	2	2	2	1	0,6	2	309482910,00
NC01	3	3	1	1	0,85	1	145683713,00
NC02	2	1	2	1	0,85	3	184945302,00
NC03	2	2	1	3	0,85	2	99073968,00
NC04	2	3	1	3	0,6	3	267309347,00
NC05	3	2	2	1	0,6	2	48762549,00
NC06	2	2	1	3	0,6	2	203627803,00
NC07	1	2	1	2	0,85	3	231297397,00
NC08	1	3	2	1	0,6	3	130318712,00
NC09	3	3	1	1	0,85	2	118597910,00
NC10	2	3	2	1	0,6	1	293764626,00
NC11	1	1	2	1	0,85	3	266835902,00
NC12	3	3	2	2	0,6	1	121904088,00
NC13	1	3	1	1	0,85	3	198628943,00
NC14	1	2	1	1	0,85	3	182423296,00
NC15	2	3	1	1	0,85	2	160873561,00
NC16	2	3	2	1	0,85	2	256640214,00
NC17	3	3	1	2	0,85	2	102624691,00
NC18	1	2	1	1	0,6	3	116997100,00
NC19	2	3	1	3	0,85	3	66784992,00
NC20	2	3	1	2	0,85	2	90676640,00
NC21	2	2	1	3	0,6	1	195795865,00
NC22	3	3	1	1	0,85	1	155876074,00
NC23	3	3	1	1	0,6	1	273562126,00
NC24	3	3	2	1	0,6	2	121822732,00
NC25	2	2	2	1	0,35	3	264279479,00
NC26	3	2	2	1	0,85	1	193459473,00
NC27	2	2	1	2	0,85	2	298436703,00
NC28	3	3	1	1	0,85	2	70822994,00
NC29	1	1	1	1	0,35	3	82539961,00

ANEXO D – Dados parciais do problema do estudo de Caso II

FATURAMENTO	FUNCIONÁRIOS	DATA CONSTITUIÇÃO	CLIENTE DESDE	SEGMENTO	RISCO	RESTRIÇÃO
2.512.000	11	novembro, 2004	maio, 2005	PEQUENA EMPRESA	C	RESTRIÇÃO IMPEDITIVA
3.893.314	6	março, 1992	janeiro, 2006	PEQUENA EMPRESA	A	SEM RESTRIÇÕES
4.638.333	30	setembro, 1988	novembro, 2004	PEQUENA EMPRESA	B	RESTRIÇÃO IMPEDITIVA
13.227.379	0	dezembro, 2007	fevereiro, 2013	EMPRESA	B	RESTRIÇÃO IMPEDITIVA
672.080	0	maio, 2013	novembro, 2013	MICROEMPRESA	B	SEM RESTRIÇÕES
1.415.162	6	junho, 2012	janeiro, 2013	PEQUENA EMPRESA	E	RESTRIÇÃO IMPEDITIVA
416.889	5	fevereiro, 1975	outubro, 1976	MICROEMPRESA	A	RESTRIÇÃO FRACA/INFO
914.286	3	abril, 1973	março, 1977	MICROEMPRESA	A	RESTRIÇÃO FRACA/INFO
87.834	1	janeiro, 1986	maio, 1986	MICROEMPRESA	B	RESTRIÇÃO FRACA/INFO
216.061	3	agosto, 1985	novembro, 1989	MICROEMPRESA	A	RESTRIÇÃO FRACA/INFO
231.716	6	julho, 1989	julho, 1990	MICROEMPRESA	A	RESTRIÇÃO IMPEDITIVA
2.111.938	12	maio, 2001	setembro, 2014	PEQUENA EMPRESA	C	SEM RESTRIÇÕES
251.120	14	julho, 1991	julho, 1999	MICROEMPRESA	C	RESTRIÇÃO IMPEDITIVA
155.800	1	novembro, 2001	janeiro, 2002	MICROEMPRESA	A	SEM RESTRIÇÕES
1.602.162	8	abril, 2004	junho, 2004	PEQUENA EMPRESA	A	SEM RESTRIÇÕES
2.658.491	19	maio, 1992	abril, 2005	PEQUENA EMPRESA	A	RESTRIÇÃO FRACA/INFO
7.825.019	20	novembro, 2003	julho, 2005	EMPRESA	A	SEM RESTRIÇÕES
121.000	0	maio, 1988	outubro, 2005	MICROEMPRESA	B	SEM RESTRIÇÕES
3.020.230	17	setembro, 2004	julho, 2006	PEQUENA EMPRESA	A	SEM RESTRIÇÕES
2.028.000	25	agosto, 1993	junho, 2006	PEQUENA EMPRESA	B	SEM RESTRIÇÕES
2.810.498	3	maio, 2006	setembro, 2011	PEQUENA EMPRESA	E	RESTRIÇÃO IMPEDITIVA
2.037.535	11	setembro, 2006	março, 2006	PEQUENA EMPRESA	A	SEM RESTRIÇÕES
271.472	1	fevereiro, 2006	fevereiro, 2006	MICROEMPRESA	C	RESTRIÇÃO FRACA/INFO
615.746	4	agosto, 2001	setembro, 2006	MICROEMPRESA	B	RESTRIÇÃO FRACA/INFO
199.350	0	abril, 1990	janeiro, 2006	MICROEMPRESA	B	SEM RESTRIÇÕES
775.100	3	dezembro, 1994	maio, 2009	MICROEMPRESA	A	SEM RESTRIÇÕES
2.376.895	12	janeiro, 1986	outubro, 2006	PEQUENA EMPRESA	A	SEM RESTRIÇÕES
13.296.652	45	setembro, 2006	novembro, 2006	EMPRESA	A	RESTRIÇÃO FRACA/INFO
918.700	1	março, 2006	setembro, 2007	MICROEMPRESA	B	SEM RESTRIÇÕES
2.052.151	13	fevereiro, 2002	fevereiro, 2007	PEQUENA EMPRESA	A	RESTRIÇÃO FRACA/INFO
392.711	6	março, 2007	outubro, 2007	MICROEMPRESA	C	RESTRIÇÃO FRACA/INFO
968.852	2	abril, 2007	maio, 2007	MICROEMPRESA	B	SEM RESTRIÇÕES
635.228	3	janeiro, 2007	agosto, 2007	MICROEMPRESA	B	SEM RESTRIÇÕES
947.389	6	agosto, 1997	novembro, 2009	MICROEMPRESA	B	RESTRIÇÃO IMPEDITIVA
2.615.658	6	junho, 2004	outubro, 2009	PEQUENA EMPRESA	E	RESTRIÇÃO IMPEDITIVA
504.419	6	novembro, 2007	setembro, 2007	MICROEMPRESA	A	RESTRIÇÃO FRACA/INFO
3.236.017	14	agosto, 2003	fevereiro, 2008	PEQUENA EMPRESA	E	RESTRIÇÃO IMPEDITIVA
2.439.222	3	abril, 2006	março, 2008	PEQUENA EMPRESA	B	SEM RESTRIÇÕES
3.023.129	3	agosto, 2007	abril, 2008	PEQUENA EMPRESA	E	RESTRIÇÃO IMPEDITIVA
1.229.720	5	julho, 2007	abril, 2008	PEQUENA EMPRESA	B	RESTRIÇÃO FRACA/INFO
301.913	2	agosto, 2008	junho, 2008	MICROEMPRESA	B	SEM RESTRIÇÕES
7.630.600	18	maio, 2005	julho, 2008	PEQUENA EMPRESA	A	SEM RESTRIÇÕES
10.810.074	176	junho, 2000	janeiro, 2009	EMPRESA	C	RESTRIÇÃO IMPEDITIVA
1.289.684	4	novembro, 2008	setembro, 2008	PEQUENA EMPRESA	A	SEM RESTRIÇÕES
1.153.724	8	maio, 2004	dezembro, 2008	PEQUENA EMPRESA	B	SEM RESTRIÇÕES
7.519.929	30	fevereiro, 2008	fevereiro, 2009	EMPRESA	A	SEM RESTRIÇÕES
495.400	3	junho, 2007	junho, 2009	MICROEMPRESA	E	RESTRIÇÃO IMPEDITIVA
2.859.661	40	setembro, 1977	dezembro, 2009	PEQUENA EMPRESA	C	RESTRIÇÃO FRACA/INFO
167.162	1	agosto, 1997	junho, 2009	MICROEMPRESA	A	SEM RESTRIÇÕES
22.449.528	110	janeiro, 1985	abril, 2003	EMPRESA	B	RESTRIÇÃO FRACA/INFO
105.676	2	junho, 1987	abril, 2001	MICROEMPRESA	E	RESTRIÇÃO IMPEDITIVA
755.446	4	julho, 2004	julho, 2009	PEQUENA EMPRESA	B	SEM RESTRIÇÕES
4.962	0	novembro, 2007	agosto, 2009	MICROEMPRESA	E	RESTRIÇÃO IMPEDITIVA
176.537	1	agosto, 2009	maio, 2009	MICROEMPRESA	B	SEM RESTRIÇÕES
364.267	0	outubro, 2000	janeiro, 2009	MICROEMPRESA	B	RESTRIÇÃO FRACA/INFO
690.851	8	fevereiro, 1986	fevereiro, 2010	MICROEMPRESA	C	RESTRIÇÃO IMPEDITIVA
2.525.815	3	agosto, 2009	abril, 2010	PEQUENA EMPRESA	B	SEM RESTRIÇÕES
2.546.600	3	outubro, 2008	abril, 2010	PEQUENA EMPRESA	E	RESTRIÇÃO IMPEDITIVA
2.400.000	15	agosto, 2010	outubro, 2010	PEQUENA EMPRESA	B	RESTRIÇÃO IMPEDITIVA

ANEXO E – Matriz de avaliação inicial para o Caso I

	c_1	c_2	c_3	c_4	c_5	c_6	$b:$
a_1	2	3	2	1	0,35	2	172018877,00
a_2	2	2	1	1	0,85	1	216278053,00
a_3	3	3	3	3	0,6	1	296494327,00
a_4	3	2	3	3	0,6	3	116691864,00
a_5	1	3	2	3	0,85	2	80254837,00
a_6	2	3	1	1	0,35	3	256396295,00
a_7	1	3	2	3	0,6	2	305070741,00
a_8	2	3	1	1	0,85	2	314741806,00
a_9	1	1	1	1	0,6	1	134817027,00
a_{10}	1	2	2	2	0,6	2	112186121,00
a_{11}	1	1	1	2	0,6	3	148564320,00
a_{12}	2	2	2	1	0,85	3	30692596,00
a_{13}	2	3	3	3	0,6	1	80892834,00
a_{14}	2	2	1	1	0,85	1	271544067,00
a_{15}	3	2	2	2	0,85	2	296802038,00
a_{16}	1	2	2	1	0,6	3	296665873,00
a_{17}	2	3	2	1	0,6	1	149901060,00
a_{18}	1	1	2	1	0,6	1	270487321,00
a_{19}	1	2	2	1	0,6	2	92613326,00
a_{20}	2	2	3	3	0,85	1	249432460,00
a_{21}	1	3	1	3	0,85	2	307422357,00
a_{22}	1	2	2	2	0,85	2	295998994,00
a_{23}	1	3	2	3	0,6	2	52853394,00
a_{24}	1	2	1	2	0,85	1	54771828,00
a_{25}	1	2	1	1	0,85	3	297145718,00
a_{26}	1	2	1	1	0,85	3	75912688,00
a_{27}	1	1	1	1	0,6	3	161234532,00
a_{28}	1	1	1	1	0,85	3	118260306,00
a_{29}	1	3	1	1	0,85	3	275922980,00
a_{30}	1	3	1	2	0,6	3	202665791,00
a_{31}	3	3	1	1	0,85	1	145683713,00
a_{32}	2	1	2	1	0,85	3	184945302,00
a_{33}	2	2	1	3	0,85	2	99073968,00
a_{34}	2	3	1	3	0,6	3	267309347,00
a_{35}	3	2	2	1	0,6	2	48762549,00
a_{36}	2	2	1	3	0,6	2	203627803,00
a_{37}	1	2	1	2	0,85	3	231297397,00
a_{38}	1	3	2	1	0,6	3	130318712,00
a_{39}	3	3	1	1	0,85	2	118597910,00
a_{40}	2	3	2	1	0,6	1	293764626,00

$B = R\$ 4.101.963.667,00$

ANEXO F – Resultados das 10 repetições da aplicação do algoritmo genético para o

 K^{AG-ED} - Caso I

Repetições do algoritmo genético para determinação do vetor K^{AG-ED}		
Teste 1	Vetor K^{AG-ED} :	0,497 0,495 0,003 0,000 0,000 0,000
	Projetos Selecionados:	1 2 3 4 5 6 9 10 12 13 15 17 19 20 23 24 26 28 30 31 33 34 35 36 38 39 40
	Custo para realização:	4097074350
	Valor do Portfólio:	86,13727458
Teste 2	Vetor K^{AG-ED} :	0,498 0,498 0,002 0,000 0,000 0,000
	Projetos Selecionados:	1 2 3 4 5 6 9 10 12 13 15 17 19 20 23 24 26 28 30 31 33 34 35 36 38 39 40
	Custo para realização:	4097074350
	Valor do Portfólio:	86,21543086
Teste 3	Vetor K^{AG-ED} :	0,493 0,493 0,006 0,006 0,000 0,000
	Projetos Selecionados:	1 2 3 4 5 6 9 10 12 13 15 17 19 20 23 24 26 28 30 31 33 34 35 36 38 39 40
	Custo para realização:	4097074350
	Valor do Portfólio:	86,06963925
Teste 4	Vetor K^{AG-ED} :	0,496 0,496 0,003 0,002 0,000 0,000
	Projetos Selecionados:	1 2 3 4 5 6 9 10 12 13 15 17 19 20 23 24 26 28 30 31 33 34 35 36 38 39 40
	Custo para realização:	4097074350
	Valor do Portfólio:	86,1725175803035
Teste 5	Vetor K^{AG-ED} :	0,493 0,493 0,012 0,001 0,000 0,000
	Projetos Selecionados:	1 2 3 4 5 6 9 10 12 13 15 17 19 20 23 24 26 28 30 31 33 34 35 36 38 39 40
	Custo para realização:	4097074350
	Valor do Portfólio:	86,030030053
Teste 6	Vetor K^{AG-ED} :	0,500 0,499 0,000 0,000 0,000 0,000
	Projetos Selecionados:	1 2 3 4 5 6 8 10 12 13 15 17 19 23 24 26 30 31 32 33 34 35 36 38 39 40
	Custo para realização:	4094251665
	Valor do Portfólio:	86,23873874050
Teste 7	Vetor K^{AG-ED} :	0,491 0,488 0,009 0,009 0,000 0,000
	Projetos Selecionados:	1 2 3 4 5 6 9 10 12 13 15 17 19 20 23 24 26 28 30 31 33 34 35 36 38 39 40
	Custo para realização:	4097074350
	Valor do Portfólio:	85,94533600851
Teste 8	Vetor K^{AG-ED} :	0,495 0,494 0,006 0,003 0,000 0,000
	Projetos Selecionados:	1 2 3 4 5 6 9 10 12 13 15 17 19 20 23 24 26 28 30 31 33 34 35 36 38 39 40
	Custo para realização:	4097074350
	Valor do Portfólio:	86,096693418
Teste 9	Vetor K^{AG-ED} :	0,500 0,500 0,000 0,000 0,000 0,000
	Projetos Selecionados:	1 2 3 4 5 6 8 10 12 13 15 17 19 23 24 26 30 31 32 33 34 35 36 38 39 40
	Custo para realização:	4094251665
	Valor do Portfólio:	86,25
Teste 10	Vetor K^{AG-ED} :	0,496 0,496 0,003 0,003 0,000 0,000
	Projetos Selecionados:	1 2 3 4 5 6 9 10 12 13 15 17 19 20 23 24 26 28 30 31 33 34 35 36 38 39 40
	Custo para realização:	4097074350
	Valor do Portfólio:	86,15981966069

**ANEXO G – Resultados das 10 repetições da aplicação do algoritmo genético para o
 K^{AG} - Caso I**

Repetições do algoritmo genético para determinação do vetor K^{AG}		
Teste 1	Vetor K^{AG}:	0,000 0,784 0,000 0,008 0,000 0,206
	Projetos Selecionados:	1 2 4 5 6 7 10 11 12 13 17 19 21 23 24 26 28 29 30 31 33 34 35 36 37 38 39
	Custo para realização:	4094041667
	Valor do Portfólio:	96,949899
Teste 2	Vetor K^{AG}:	0,000 0,810 0,011 0,000 0,000 0,177
	Projetos Selecionados:	1 4 5 6 7 10 11 12 13 17 19 21 23 24 26 27 28 29 30 31 33 34 35 36 38 39 40
	Custo para realização:	4101465375
	Valor do Portfólio:	97,00701414672
Teste 3	Vetor K^{AG}:	0,000 0,839 0,006 0,002 0,000 0,151
	Projetos Selecionados:	1 3 4 5 6 7 10 12 13 17 19 21 23 24 26 28 29 30 31 33 34 35 36 38 39 40
	Custo para realização:	4088160850
	Valor do Portfólio:	97,50601205090
Teste 4	Vetor K^{AG}:	0,000 0,102 0,001 0,000 0,000 0,896
	Projetos Selecionados:	1 4 5 6 10 11 12 13 16 17 19 23 24 25 26 27 28 29 30 32 33 34 35 36 37 38 39
	Custo para realização:	4059578228
	Valor do Portfólio:	98,198198178986
Teste 5	Vetor K^{AG}:	0,000 0,069 0,000 0,000 0,000 0,930
	Projetos Selecionados:	1 4 5 6 10 11 12 13 16 17 19 23 24 25 26 27 28 29 30 32 33 34 35 36 37 38 39
	Custo para realização:	4059578228
	Valor do Portfólio:	98,48198198199
Teste 6	Vetor K^{AG}:	0,000 0,782 0,000 0,000 0,009 0,207
	Projetos Selecionados:	1 2 4 5 6 7 8 10 11 12 13 17 19 23 24 26 28 29 30 31 33 34 35 36 37 38 39
	Custo para realização:	4101361116
	Valor do Portfólio:	96,449417512995
Teste 7	Vetor K^{AG}:	0,000 0,169 0,000 0,000 0,000 0,830
	Projetos Selecionados:	1 4 5 6 10 11 12 13 16 17 19 23 24 25 26 27 28 29 30 32 33 34 35 36 37 38 39
	Custo para realização:	4059578228
	Valor do Portfólio:	97,73123124637
Teste 8	Vetor K^{AG}:	0,000 0,102 0,001 0,000 0,000 0,896
	Projetos Selecionados:	1 4 5 6 10 11 12 13 16 17 19 23 24 25 26 27 28 29 30 32 33 34 35 36 37 38 39
	Custo para realização:	4059578228
	Valor do Portfólio:	98,19819817898
Teste 9	Vetor K^{AG}:	0,000 0,847 0,027 0,000 0,000 0,125
	Projetos Selecionados:	1 3 4 5 6 7 8 10 12 13 17 19 23 24 26 28 29 30 31 33 34 35 36 38 39 40
	Custo para realização:	4095480299
	Valor do Portfólio:	97,2747747803
Teste 10	Vetor K^{AG}:	0,010 0,802 0,000 0,052 0,000 0,135
	Projetos Selecionados:	1 3 4 5 6 7 10 12 13 17 19 21 23 24 26 28 29 30 31 33 34 35 36 38 39 40
	Custo para realização:	4088160850
	Valor do Portfólio:	96,33183172420

ANEXO H – Matriz de avaliação inicial para o Caso II

	c_1	c_2	c_3	c_4	c_5	c_6	c_7	$b:$
a_1	2	3	2	2	1	2	2	5000
a_2	1	1	1	2	3	2	2	5300
a_3	2	2	2	4	3	3	1	27118
a_4	1	1	4	3	3	2	2	463
a_5	3	3	4	4	1	2	3	15000
a_6	3	3	4	4	1	2	3	50000
a_7	2	2	2	1	1	2	3	49581
a_8	2	2	2	1	1	2	3	45097
a_9	3	4	1	1	1	2	3	110000
a_{10}	2	2	2	1	3	2	2	5000
a_{11}	4	4	3	3	1	1	2	10000
a_{12}	3	3	4	4	1	2	2	20000
a_{13}	4	4	2	3	2	2	1	6500
a_{14}	4	3	1	1	2	2	3	10000
a_{15}	4	3	1	1	2	2	2	24900
a_{16}	4	3	1	1	2	2	2	24900
a_{17}	4	4	3	2	2	5	1	30470
a_{18}	4	2	2	1	1	2	2	60000
a_{19}	4	2	2	1	1	2	2	50000
a_{20}	3	1	4	4	1	2	2	55500
a_{21}	3	1	4	4	1	2	2	69500
a_{22}	4	4	3	3	4	3	2	200000
a_{23}	4	4	3	3	4	3	2	395000
a_{24}	2	2	3	3	1	5	1	44107
a_{25}	3	2	4	4	1	2	1	51300
a_{26}	1	1	2	1	3	2	3	6000
a_{27}	1	1	2	1	3	2	3	10000
a_{28}	1	1	2	1	3	2	3	8900
a_{29}	2	3	2	2	3	2	1	10000
a_{30}	3	1	3	3	1	5	1	39659
a_{31}	2	1	1	1	3	3	2	1591
a_{32}	3	1	1	1	1	2	3	30000
a_{33}	4	3	2	1	2	2	2	25232
a_{34}	1	1	4	4	3	4	1	371
a_{35}	1	3	1	1	3	2	3	5000
a_{36}	3	2	2	1	1	1	2	10000
a_{37}	3	3	2	2	1	2	1	50000
a_{38}	1	3	2	2	3	2	2	11700
a_{39}	4	4	1	1	2	1	3	59300
a_{40}	1	1	1	2	3	2	2	979
								$B = R\\$ 898.410,00$

ANEXO I – Matriz de registros históricos (parcial) do Caso II

	c_1	c_2	c_3	c_4	c_5	c_6	c_7	y^h :
a_1^h	3	3	2	3	2	3	1	Adimplente
a_2^h	4	3	4	3	2	5	3	Adimplente
a_3^h	4	4	4	3	2	4	1	Adimplente
a_4^h	4	1	2	1	3	4	1	Inadimplente
a_5^h	2	1	1	1	1	4	3	Adimplente
a_6^h	3	3	1	1	2	1	1	Inadimplente
a_7^h	1	3	4	4	1	5	2	Adimplente
a_8^h	2	2	4	4	1	5	2	Adimplente
a_9^h	1	1	4	4	1	4	2	Inadimplente
a_{10}^h	1	2	4	4	1	5	2	Inadimplente
a_{11}^h	1	3	4	4	1	5	1	Adimplente
a_{12}^h	3	4	3	1	2	3	3	Adimplente
a_{13}^h	1	4	4	4	1	3	1	Inadimplente
a_{14}^h	1	1	3	4	1	5	3	Adimplente
a_{15}^h	3	3	2	3	2	5	3	Adimplente
a_{16}^h	3	4	4	3	2	5	2	Adimplente
a_{17}^h	4	4	3	3	3	5	3	Adimplente
a_{18}^h	1	1	4	3	1	4	3	Adimplente
a_{19}^h	3	4	2	3	2	5	3	Adimplente
a_{20}^h	3	4	4	3	2	4	3	Adimplente
a_{21}^h	3	2	2	1	2	1	1	Inadimplente
a_{22}^h	3	3	2	3	2	5	3	Adimplente
a_{23}^h	1	1	2	3	1	3	2	Adimplente
a_{24}^h	2	2	3	3	1	4	2	Adimplente
a_{25}^h	1	1	4	3	1	4	3	Adimplente
a_{26}^h	2	2	4	2	1	5	3	Adimplente
a_{27}^h	3	4	4	3	2	5	3	Inadimplente
a_{28}^h	4	4	2	3	3	5	2	Adimplente
a_{29}^h	2	1	2	3	1	4	3	Adimplente
a_{30}^h	3	4	3	3	2	5	2	Adimplente
a_{31}^h	1	3	2	3	1	3	2	Adimplente
a_{32}^h	2	2	2	3	1	4	3	Adimplente
a_{33}^h	2	2	2	3	1	4	3	Adimplente
a_{34}^h	2	3	3	2	1	4	1	Adimplente
a_{35}^h	3	3	2	2	2	1	1	Inadimplente
a_{36}^h	1	3	2	3	1	5	2	Adimplente
a_{37}^h	4	4	3	3	2	1	1	Inadimplente
a_{38}^h	3	2	2	2	2	4	3	Adimplente
...
a_{5432}^h	1	2	3	3	1	4	3	Adimplente

ANEXO J – Matriz de avaliação modificada ($- a_i | y_i = 0$) para o Caso II

	c_1	c_2	c_3	c_4	c_5	c_6	c_7	$b:$
a_1	2	3	2	2	1	2	2	5000
a_2	1	1	1	2	3	2	2	5300
a_4	1	1	4	3	3	2	2	463
a_5	3	3	4	4	1	2	3	15000
a_6	3	3	4	4	1	2	3	50000
a_7	2	2	2	1	1	2	3	49581
a_8	2	2	2	1	1	2	3	45097
a_9	3	4	1	1	1	2	3	110000
a_{10}	2	2	2	1	3	2	2	5000
a_{12}	3	3	4	4	1	2	2	20000
a_{14}	4	3	1	1	2	2	3	10000
a_{15}	4	3	1	1	2	2	2	24900
a_{16}	4	3	1	1	2	2	2	24900
a_{17}	4	4	3	2	2	5	1	30470
a_{18}	4	2	2	1	1	2	2	60000
a_{19}	4	2	2	1	1	2	2	50000
a_{20}	3	1	4	4	1	2	2	55500
a_{21}	3	1	4	4	1	2	2	69500
a_{22}	4	4	3	3	4	3	2	200000
a_{23}	4	4	3	3	4	3	2	395000
a_{24}	2	2	3	3	1	5	1	44107
a_{26}	1	1	2	1	3	2	3	6000
a_{27}	1	1	2	1	3	2	3	10000
a_{28}	1	1	2	1	3	2	3	8900
a_{30}	3	1	3	3	1	5	1	39659
a_{31}	2	1	1	1	3	3	2	1591
a_{32}	3	1	1	1	1	2	3	30000
a_{33}	4	3	2	1	2	2	2	25232
a_{35}	1	3	1	1	3	2	3	5000
a_{38}	1	3	2	2	3	2	2	11700
a_{40}	1	1	1	2	3	2	2	979
								$B = R\\$ 898.410,00$

ANEXO K – Matriz de avaliação modificada ($+c_{m+1}$) para o Caso II

	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	$b:$
a_1	2	3	2	2	1	2	2	1	5000
a_2	1	1	1	2	3	2	2	1	5300
a_3	2	2	2	4	3	3	1	0	27118
a_4	1	1	4	3	3	2	2	1	463
a_5	3	3	4	4	1	2	3	1	15000
a_6	3	3	4	4	1	2	3	1	50000
a_7	2	2	2	1	1	2	3	1	49581
a_8	2	2	2	1	1	2	3	1	45097
a_9	3	4	1	1	1	2	3	1	110000
a_{10}	2	2	2	1	3	2	2	1	5000
a_{11}	4	4	3	3	1	1	2	0	10000
a_{12}	3	3	4	4	1	2	2	1	20000
a_{13}	4	4	2	3	2	2	1	0	6500
a_{14}	4	3	1	1	2	2	3	1	10000
a_{15}	4	3	1	1	2	2	2	1	24900
a_{16}	4	3	1	1	2	2	2	1	24900
a_{17}	4	4	3	2	2	5	1	1	30470
a_{18}	4	2	2	1	1	2	2	1	60000
a_{19}	4	2	2	1	1	2	2	1	50000
a_{20}	3	1	4	4	1	2	2	1	55500
a_{21}	3	1	4	4	1	2	2	1	69500
a_{22}	4	4	3	3	4	3	2	1	200000
a_{23}	4	4	3	3	4	3	2	1	395000
a_{24}	2	2	3	3	1	5	1	1	44107
a_{25}	3	2	4	4	1	2	1	0	51300
a_{26}	1	1	2	1	3	2	3	1	6000
a_{27}	1	1	2	1	3	2	3	1	10000
a_{28}	1	1	2	1	3	2	3	1	8900
a_{29}	2	3	2	2	3	2	1	0	10000
a_{30}	3	1	3	3	1	5	1	1	39659
a_{31}	2	1	1	1	3	3	2	1	1591
a_{32}	3	1	1	1	1	2	3	1	30000
a_{33}	4	3	2	1	2	2	2	1	25232
a_{34}	1	1	4	4	3	4	1	0	371
a_{35}	1	3	1	1	3	2	3	1	5000
a_{36}	3	2	2	1	1	1	2	0	10000
a_{37}	3	3	2	2	1	2	1	0	50000
a_{38}	1	3	2	2	3	2	2	1	11700
a_{39}	4	4	1	1	2	1	3	0	59300
a_{40}	1	1	1	2	3	2	2	1	979
									$B = R\\$ 898.410,00$

**ANEXO L – Resultados das 10 repetições da aplicação do algoritmo genético
(MA e K^{AG-FS} - Caso II)**

Repetições do algoritmo genético para determinação do vetor K^{AG-FS}		
Teste 1	Vetor K^{AG-FS}:	0,127 0,000 0,128 0,128 0,127 0,245 0,243
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 24 25 26 27 28 29 30 31 32 33 34 35 36 38 39 40
	Custo para realização:	878468
	Valor do Portfólio:	107,294785
Teste 2	Vetor K^{AG-FS}:	0,147 0,000 0,152 0,154 0,148 0,199 0,197
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 24 25 26 27 28 29 30 31 32 33 34 35 36 38 39 40
	Custo para realização:	878468
	Valor do Portfólio:	107,145735
Teste 3	Vetor K^{AG-FS}:	0,012 0,001 0,017 0,020 0,016 0,475 0,457
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 24 25 26 27 28 29 30 31 32 33 34 35 36 38 39 40
	Custo para realização:	878468
	Valor do Portfólio:	107,701276
Teste 4	Vetor K^{AG-FS}:	0,112 0,000 0,114 0,114 0,112 0,274 0,272
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 24 25 26 27 28 29 30 31 32 33 34 35 36 38 39 40
	Custo para realização:	878468
	Valor do Portfólio:	107,368252
Teste 5	Vetor K^{AG-FS}:	0,119 0,002 0,129 0,129 0,119 0,251 0,246
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 24 25 26 27 28 29 30 31 32 33 34 35 36 38 39 40
	Custo para realização:	878468
	Valor do Portfólio:	107,278418
Teste 6	Vetor K^{AG-FS}:	0,126 0,000 0,130 0,131 0,126 0,244 0,239
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 24 25 26 27 28 29 30 31 32 33 34 35 36 38 39 40
	Custo para realização:	878468
	Valor do Portfólio:	107,264956
Teste 7	Vetor K^{AG-FS}:	0,124 0,003 0,133 0,133 0,124 0,241 0,239
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 24 25 26 27 28 29 30 31 32 33 34 35 36 38 39 40
	Custo para realização:	878468
	Valor do Portfólio:	107,270565
Teste 8	Vetor K^{AG-FS}:	0,135 0,008 0,136 0,136 0,135 0,224 0,223
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 24 25 26 27 28 29 30 31 32 33 34 35 36 38 39 40
	Custo para realização:	878468
	Valor do Portfólio:	107,236904
Teste 9	Vetor K^{AG-FS}:	0,129 0,020 0,133 0,133 0,129 0,226 0,226
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 24 25 26 27 28 29 30 31 32 33 34 35 36 38 39 40
	Custo para realização:	878468
	Valor do Portfólio:	107,229283
Teste 10	Vetor K^{AG-FS}:	0,123 0,003 0,125 0,125 0,123 0,250 0,248
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 24 25 26 27 28 29 30 31 32 33 34 35 36 38 39 40
	Custo para realização:	878468
	Valor do Portfólio:	107,303063

**ANEXO M – Resultados das 10 repetições da aplicação do algoritmo genético
(MA e K^{AG-ED} - Caso II)**

Repetições do algoritmo genético para determinação do vetor K^{AG-ED}		
Teste 1	Vetor K^{AG-ED} :	0,493 0,000 0,000 0,001 0,000 0,004 0,499
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	117,792518
Teste 2	Vetor K^{AG-ED} :	0,498 0,000 0,000 0,000 0,000 0,001 0,499
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	117,889081
Teste 3	Vetor K^{AG-ED} :	0,492 0,000 0,000 0,002 0,000 0,008 0,496
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	117,719326
Teste 4	Vetor K^{AG-ED} :	0,492 0,000 0,000 0,000 0,000 0,011 0,495
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	117,707958
Teste 5	Vetor K^{AG-ED} :	0,496 0,000 0,000 0,000 0,000 0,005 0,498
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	117,815627
Teste 6	Vetor K^{AG-ED} :	0,495 0,000 0,000 0,000 0,000 0,002 0,501
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	117,843441
Teste 7	Vetor K^{AG-ED} :	0,492 0,000 0,000 0,000 0,000 0,015 0,492
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	117,657451
Teste 8	Vetor K^{AG-ED} :	0,493 0,000 0,000 0,000 0,000 0,006 0,500
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	117,770027
Teste 9	Vetor K^{AG-ED} :	0,493 0,000 0,000 0,000 0,000 0,007 0,499
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	117,758808
Teste 10	Vetor K^{AG-ED} :	0,495 0,000 0,000 0,002 0,000 0,003 0,497
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	117,815544

**ANEXO N – Resultados das 10 repetições da aplicação do algoritmo genético
(MA e K^{AG} - Caso II)**

Repetições do algoritmo genético para determinação do vetor K^{AG}		
Teste 1	Vetor K^{AG}:	0,748 0,003 0,000 0,000 0,000 0,000 0,248
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	120,834467
Teste 2	Vetor K^{AG}:	0,849 0,000 0,028 0,000 0,000 0,013 0,108
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	121,906723
Teste 3	Vetor K^{AG}:	0,829 0,000 0,000 0,000 0,017 0,000 0,153
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	121,521487
Teste 4	Vetor K^{AG}:	0,769 0,000 0,038 0,022 0,019 0,004 0,144
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	120,3985018
Teste 5	Vetor K^{AG}:	0,784 0,007 0,000 0,003 0,011 0,000 0,191
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	121,042388
Teste 6	Vetor K^{AG}:	0,835 0,000 0,000 0,005 0,000 0,000 0,159
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	121,863270
Teste 7	Vetor K^{AG}:	0,880 0,023 0,031 0,000 0,000 0,000 0,065
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	122,317717
Teste 8	Vetor K^{AG}:	0,765 0,000 0,000 0,000 0,000 0,000 0,234
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	121,055800
Teste 9	Vetor K^{AG}:	0,720 0,000 0,090 0,000 0,006 0,000 0,181
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	120,220487
Teste 10	Vetor K^{AG}:	0,890 0,000 0,055 0,000 0,000 0,009 0,044
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	122,415692

ANEXO O – Resultados das 10 repetições da aplicação do algoritmo genético

 $(MA_{(-a_i|y_i=0)} \text{ e } K^{AG-FS} - \text{Caso II})$

Repetições do algoritmo genético para determinação do vetor K^{AG-FS}		
Teste 1	Vetor K^{AG-FS} :	0,002 0,000 0,020 0,024 0,007 0,822 0,120
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	110,224322
Teste 2	Vetor K^{AG-FS} :	0,015 0,000 0,048 0,049 0,016 0,730 0,140
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	107,921951
Teste 3	Vetor K^{AG-FS} :	0,009 0,000 0,036 0,041 0,020 0,795 0,094
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	109,064217
Teste 4	Vetor K^{AG-FS} :	0,010 0,000 0,025 0,029 0,017 0,800 0,117
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	109,521249
Teste 5	Vetor K^{AG-FS} :	0,030 0,000 0,041 0,046 0,032 0,735 0,114
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	107,757314
Teste 6	Vetor K^{AG-FS} :	0,015 0,000 0,036 0,044 0,030 0,740 0,133
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	107,987806
Teste 7	Vetor K^{AG-FS} :	0,002 0,000 0,030 0,061 0,022 0,806 0,076
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	108,822155
Teste 8	Vetor K^{AG-FS} :	0,000 0,000 0,030 0,051 0,017 0,792 0,108
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	108,931374
Teste 9	Vetor K^{AG-FS} :	0,001 0,000 0,011 0,016 0,007 0,831 0,130
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	110,586677
Teste 10	Vetor K^{AG-FS} :	0,001 0,000 0,014 0,021 0,001 0,851 0,107
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	110,863443

ANEXO P – Resultados das 10 repetições da aplicação do algoritmo genético

 $(MA_{(-a_i|y_i=0)} e K^{AG-ED} - \text{Caso II})$

Repetições do algoritmo genético para determinação do vetor K^{AG-ED}		
Teste 1	Vetor K^{AG-ED} :	0,332 0,000 0,000 0,000 0,000 0,332 0,334
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	103,410895
Teste 2	Vetor K^{AG-ED} :	0,333 0,000 0,000 0,000 0,000 0,332 0,334
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	103,404300
Teste 3	Vetor K^{AG-ED} :	0,332 0,000 0,000 0,000 0,000 0,332 0,335
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	103,406480
Teste 4	Vetor K^{AG-ED} :	0,333 0,000 0,000 0,000 0,000 0,332 0,334
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	103,404300
Teste 5	Vetor K^{AG-ED} :	0,333 0,000 0,000 0,000 0,000 0,332 0,334
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	103,404300
Teste 6	Vetor K^{AG-ED} :	0,333 0,000 0,000 0,000 0,000 0,332 0,334
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	103,40430
Teste 7	Vetor K^{AG-ED} :	0,332 0,000 0,000 0,000 0,000 0,332 0,335
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	103,406480
Teste 8	Vetor K^{AG-ED} :	0,332 0,000 0,000 0,000 0,000 0,332 0,334
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	103,410895
Teste 9	Vetor K^{AG-ED} :	0,331 0,000 0,000 0,004 0,000 0,329 0,334
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	103,27710
Teste 10	Vetor K^{AG-ED} :	0,332 0,000 0,000 0,000 0,000 0,332 0,334
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	103,410895

ANEXO Q – Resultados das 10 repetições da aplicação do algoritmo genético

(MA_(-a_i|y_i=0) e K^{AG} - Caso II)

Repetições do algoritmo genético para determinação do vetor K ^{AG}		
Teste 1	Vetor K ^{AG} :	0,039 0,010 0,000 0,005 0,000 0,928 0,017
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	112,141136
Teste 2	Vetor K ^{AG} :	0,013 0,001 0,010 0,008 0,000 0,893 0,072
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	111,691408
Teste 3	Vetor K ^{AG} :	0,019 0,000 0,000 0,000 0,015 0,817 0,148
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	110,602957
Teste 4	Vetor K ^{AG} :	0,000 0,002 0,009 0,005 0,000 0,914 0,069
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	112,015806
Teste 5	Vetor K ^{AG} :	0,000 0,043 0,000 0,000 0,006 0,842 0,107
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	110,592154
Teste 6	Vetor K ^{AG} :	0,000 0,002 0,000 0,012 0,003 0,877 0,104
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	111,427214
Teste 7	Vetor K ^{AG} :	0,000 0,043 0,000 0,000 0,000 0,852 0,103
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	110,851970
Teste 8	Vetor K ^{AG} :	0,024 0,000 0,010 0,000 0,000 0,917 0,048
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	112,142323
Teste 9	Vetor K ^{AG} :	0,007 0,000 0,010 0,000 0,000 0,785 0,197
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	110,400277
Teste 10	Vetor K ^{AG} :	0,000 0,000 0,001 0,005 0,022 0,899 0,071
	Projetos Selecionados:	1 2 4 5 6 7 8 9 10 12 14 15 16 17 18 19 20 21 24 26 27 28 30 31 32 33 35 38 40
	Custo para realização:	813879
	Valor do Portfólio:	111,4803248

ANEXO R – Resultados das 10 repetições da aplicação do algoritmo genético

 $(MA_{(+c_{m+1})}$ e K^{AG-ED} - Caso II)

Repetições do algoritmo genético para determinação do vetor K^{AG-ED}		
Teste 1	Vetor K^{AG-ED} :	0,486 0,000 0,000 0,000 0,000 0,022 0,491 0,000
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	117,509827
Teste 2	Vetor K^{AG-ED} :	0,490 0,000 0,000 0,000 0,000 0,016 0,493 0,000
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	117,623275
Teste 3	Vetor K^{AG-ED} :	0,481 0,000 0,000 0,000 0,000 0,033 0,484 0,000
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	117,33332
Teste 4	Vetor K^{AG-ED} :	0,493 0,000 0,000 0,000 0,000 0,008 0,498 0,000
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	117,7475854
Teste 5	Vetor K^{AG-ED} :	0,490 0,000 0,000 0,002 0,001 0,006 0,493 0,003
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	117,540059
Teste 6	Vetor K^{AG-ED} :	0,499 0,000 0,000 0,000 0,000 0,000 0,500 0,000
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	117,90603
Teste 7	Vetor K^{AG-ED} :	0,495 0,000 0,000 0,000 0,000 0,004 0,498 0,000
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	117,826665
Teste 8	Vetor K^{AG-ED} :	0,493 0,000 0,000 0,000 0,000 0,009 0,497 0,000
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	117,736358
Teste 9	Vetor K^{AG-ED} :	0,493 0,000 0,000 0,000 0,000 0,004 0,499 0,001
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	117,7450135
Teste 10	Vetor K^{AG-ED} :	0,497 0,000 0,000 0,000 0,000 0,003 0,498 0,000
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	117,855183

ANEXO S – Resultados das 10 repetições da aplicação do algoritmo genético

(MA_(+ c_{m+1}) e K^{AG} - Caso II)

Repetições do algoritmo genético para determinação do vetor K ^{AG}		
Teste 1	Vetor K ^{AG} :	0,931 0,009 0,000 0,000 0,000 0,036 0,022 0,000
	Projetos Selecionados:	1 2 3 4 5 6 7 10 11 12 13 14 15 16 17 18 19 20 21 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	883371
	Valor do Portfólio:	122,757638
Teste 2	Vetor K ^{AG} :	0,917 0,025 0,000 0,004 0,000 0,022 0,031 0,000
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	122,550998
Teste 3	Vetor K ^{AG} :	0,870 0,005 0,000 0,000 0,000 0,007 0,117 0,000
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	122,256931
Teste 4	Vetor K ^{AG} :	0,770 0,014 0,010 0,000 0,000 0,065 0,138 0,000
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	120,324070
Teste 5	Vetor K ^{AG} :	0,869 0,001 0,020 0,000 0,000 0,016 0,090 0,000
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	122,157724
Teste 6	Vetor K ^{AG} :	0,946 0,000 0,002 0,000 0,000 0,000 0,048 0,000
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	123,350967
Teste 7	Vetor K ^{AG} :	0,849 0,000 0,042 0,000 0,000 0,041 0,066 0,000
	Projetos Selecionados:	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	884361
	Valor do Portfólio:	121,568546
Teste 8	Vetor K ^{AG} :	0,917 0,000 0,000 0,008 0,000 0,038 0,036 0,000
	Projetos Selecionados:	1 2 3 4 5 6 8 10 11 12 13 14 15 16 17 18 19 20 21 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	878887
	Valor do Portfólio:	122,477860
Teste 9	Vetor K ^{AG} :	0,901 0,036 0,000 0,002 0,000 0,038 0,021 0,000
	Projetos Selecionados:	1 2 3 4 5 6 7 10 11 12 13 14 15 16 17 18 19 20 21 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	883371
	Valor do Portfólio:	122,221723
Teste 10	Vetor K ^{AG} :	0,945 0,011 0,039 0,000 0,000 0,003 0,000 0,000
	Projetos Selecionados:	1 2 3 4 5 6 7 10 11 12 13 14 15 16 17 18 19 20 21 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
	Custo para realização:	883371
	Valor do Portfólio:	123,227176