



PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PROCESSOS
MESTRADO PROFISSIONAL

MODELO FUZZY PARA AVALIAÇÃO DE IMÓVEIS UTILIZANDO ÁRVORE DE DECISÃO

Jairo Fadul de Lima

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia de Processos – Mestrado Profissional, PPGEP/ITEC, da Universidade Federal do Pará, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia de Processos.

Orientadores: Carlos Tavares Costa Júnior
Emanuel Negrão Macêdo

Belém
Dezembro de 2017

**MODELO FUZZY PARA AVALIAÇÃO DE IMÓVEIS
UTILIZANDO ÁRVORE DE DECISÃO**

Jairo Fadul de Lima

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO PROGRAMA DE PÓSGRADUAÇÃO EM ENGENHARIA PROCESSOS – MESTRADO PROFISSIONAL (PPGEP/ITEC) DA UNIVERSIDADE FEDERAL DO PARÁ COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM ENGENHARIA DE PROCESSOS.

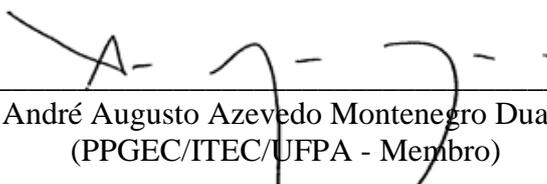
Examinada por:



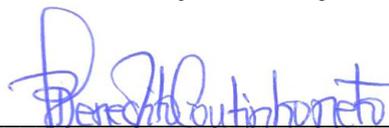
Prof. Carlos Tavares Costa Júnior, Dr.
(PPGEP/ITEC/UFPA - Orientador)



Prof. Emanuel Negrão Macêdo, D. Sc.
(PPGEP/ITEC/UFPA - Coorientador)



Prof. André Augusto Azevedo Montenegro Duarte, Dr.
(PPGEP/ITEC/UFPA - Membro)



Prof. Benedito Coutinho Neto, Dr.
(IFPA – Membro)

BELÉM, PA - BRASIL

DEZEMBRO DE 2017

**Dados Internacionais de Catalogação-na-Publicação (CIP)
Sistema de Bibliotecas da UFPA**

Lima, Jairo Fadul de, 1956-
Modelo Fuzzy para avaliação de imóveis utilizando árvore
de decisão/ Jairo Fadul de Lima.- 2017.

Orientadores: Carlos Tavares Costa Júnior
Emanuel Negrão Macêdo

Dissertação (Mestrado Profissional) - Universidade
Federal do Pará. Instituto de Tecnologia. Programa de Pós-
Graduação em Engenharia de Processos, 2017

1. Bens imóveis- Avaliação- Modelos matemáticos 2. Lógica
difusa 3. Computação- Matemática I. Título

CDD 22.ed.333.332

Resumo da Dissertação apresentada ao PPGEP/UFPA como parte dos requisitos necessários para a obtenção do grau de Mestre em Engenharia de Processos (M. Eng.)

MODELO FUZZY PARA AVALIAÇÃO DE IMÓVEIS UTILIZANDO ÁRVORE DE DECISÃO

Jairo Fadul de Lima

Dezembro/2017

Orientadores: Carlos Tavares Costa Júnior
Emanuel Negrão Macêdo

Área de Concentração: Engenharia de Processos

Este trabalho tem como objetivo propor um modelo de avaliação de imóveis, que incorpore as imprecisões e subjetividades inerentes às variáveis que compõem o mercado imobiliário, para ser utilizado por profissionais que atuam na área. O modelo desenvolvido utiliza a lógica Fuzzy, para definir as características subjetivas dos imóveis, na forma de variáveis linguísticas, atribuindo a cada uma delas funções de pertinência, refletindo a forma de pensar dos especialistas humanos em avaliações imobiliárias. A pesquisa aborda também o uso de técnicas de extração de regras de classificação de atributos relevantes, chamada de Árvore de Decisão, para a redução das regras originalmente estabelecidas no modelo Fuzzy proposto, sem perder o foco na lógica do processo. Esta redução objetiva viabilizar o uso prático do modelo proposto. Para validação é feita uma análise comparativa da utilização do modelo Fuzzy, ao qual foram submetidos dois grupos e os valores gerados foram comparados com dados reais de mercado, obtendo-se os Desvios Padrão na faixa de R\$ 600,00 e os percentuais de desvios das médias dos valores calculados e levantados em campo, em relação ao Desvio Padrão, na faixa de 15% e 17%, respectivamente, para ambos os grupos, mostrando uma uniformidade nos cálculos executados pelo modelo Fuzzy proposto.

Abstract of Dissertation presented to PPGE/UFPA as a partial fulfillment of the requirements for the degree of Master in Process Engineering (M. Eng.)

**FUZZY MODEL FOR ASSESSMENT OF BUILDINGS
USING DECISION TREE**

Jairo Fadul de Lima

December/2017

Advisors: Carlos Tavares Costa Júnior
Emanuel Negrão Macêdo

Research Area: Process Engineering

This paper aims to propose a real estate appraisal model that incorporates the imprecisions and subjectivities inherent to the variables that make up the real estate market, to be used by professionals working in the area. The developed model uses Fuzzy logic to define the subjective characteristics of real estate, in the form of linguistic variables, assigning to each of them pertinence functions, reflecting the thinking of human specialists in real estate appraisals. The research also addresses the use of rules extraction techniques for the classification of relevant attributes, called the Decision Tree, to reduce the rules originally established in the proposed Fuzzy model, without losing focus on the logic of the process. This objective reduction makes feasible the practical use of the proposed model. To validate, a comparative analysis of the use of the Fuzzy model was performed, to which two groups were submitted and the generated values were compared with real market data, obtaining the Standard Deviations in the range of R \$ 600.00 and the percentages of deviations of the values of the values generated and collected in the field, in relation to the Standard Deviation in the range of 15% and 17%, for both groups, showing a uniformity in the calculations performed by the proposed Fuzzy model.

SUMÁRIO

CAPÍTULO 1 - INTRODUÇÃO.....	1
1.1 - MOTIVAÇÃO.....	1
1.2 - OBJETIVOS.....	1
1.2.1 - Objetivo geral.....	1
1.2.2 - Objetivos específicos.....	2
1.3 – CONTRIBUIÇÕES DA DISSERTAÇÃO.....	2
1.4 - ORGANIZAÇÃO DO TRABALHO.....	5
CAPÍTULO 2 – REVISÃO DA LITERATURA.....	7
2.1 – AVALIAÇÃO IMOBILIÁRIA.....	7
2.2 – MÉTODOS DE AVALIAÇÃO DE IMÓVEIS.....	8
2.2.1 – Método comparativo de dados de mercado.....	8
2.2.2 – Método comparativo do custo de reprodução.....	9
2.2.3 – Método da renda.....	9
2.2.4 – Método involutivo.....	10
2.2.5 – Método residual.....	11
2.2.6 – Modelo por regressão.....	11
2.3 – LÓGICA FUZZY.....	13
2.4 – FUNDAMENTOS DA LÓGICA FUZZY.....	14
2.4.1 - Conjuntos Fuzzy.....	17
2.4.2 - Variáveis linguísticas e modificadores.....	18
2.4.3 - Base de regras.....	18
2.4.4 – Inferência Fuzzy.....	19
2.4.4.1 - Estilo Mamdani.....	19
2.4.4.2 - Estilo Sugeno.....	20
2.4.4.3 – Fuzzyficação.....	21
2.4.5 – Avaliação das regras Fuzzy.....	23
2.4.6 – Agregação das regras Fuzzy.....	23
2.4.7 – Defuzzyficação.....	24
2.5 – ÁRVORE DE DECISÃO.....	25
2.6 – CLASSIFICAÇÃO EM ÁRVORE DE DECISÃO.....	27
2.7 – TÉCNICAS PARA A SELEÇÃO DE ATRIBUTOS.....	29

2.7.1 - Exemplo de indução de árvores de decisão.....	29
2.8 – ALGORITMO ID3.....	32
2.9 – ALGORITMO C4.5.....	35
2.10 – ALGORITMO J48.....	36
2.11 – SOFTWARE WEKA.....	37
2.11.1 – Árvore de decisão.....	37
CAPÍTULO 3 - METODOLOGIA EXPERIMENTAL.....	43
3.1 – METODOLOGIA EXPERIMENTAL.....	43
3.2 - CRIAÇÃO DO MODELO FUZZY.....	44
3.3 – CRIAÇÃO DA ÁRVORE DE DECISÃO.....	51
3.3.1 - Utilização da Ferramenta WEKA.....	53
3.4 - NOVO CONJUNTO DE REGRAS.....	56
CAPÍTULO 4 - RESULTADOS E DISCUSSÃO.....	57
4.1 - APLICANDO OS DADOS.....	57
4.1.1 - Processamento do primeiro grupo de dados.....	57
4.1.2 - Processamento do segundo grupo de dados.....	60
4.2 – DISCUSSÃO DOS RESULTADOS.....	62
CAPÍTULO 5 - CONCLUSÕES E SUGESTÕES.....	63
5.1 - CONCLUSÕES.....	63
5.2 - SUGESTÕES.....	64
REFERÊNCIAS BIBLIOGRÁFICAS.....	65
APÊNDICE A - CONJUNTO DE TREINO/TESTE UTILIZADO PARA A GERAÇÃO DA ÁRVORE DE DECISÃO.....	69
APÊNDICE B - RESULTADOS OBTIDOS PELO MODELO FUZZY PROPOSTO.....	71
APÊNDICE C - ÁRVORE DE DECISÃO GERADA PARA SIMPLIFICAR AS REGRAS DO MODELO.....	72
ANEXO I - LEVANTAMENTO DE CAMPO FEITO PELA FACULDADE DE ENGENHARIA CIVIL/ITEC/UFPA.....	73
ANEXO II - IMÓVEIS INFERIDOS NO MODELO.....	77

LISTA DE FIGURAS

Figura 2.1	Reta de regressão.....	12
Figura 2.2	Representação na forma de conjuntos da altura de uma pessoa, sob o ponto de vista da Lógica convencional (à esquerda) e do da Lógica Fuzzy (à direita).....	15
Figura 2.3	Tipos de função de pertinência.....	17
Figura 2.4	Sistema Fuzzy de Mamdani.....	20
Figura 2.5	Sistema Fuzzy de Sugeno.....	21
Figura 2.6	Funções de pertinência para o exemplo da gorjeta.....	22
Figura 2.7	Grau de pertinência para o exemplo da gorjeta.....	22
Figura 2.8	Conjuntos Fuzzy produzidos pela variável linguística gorjeta.....	24
Figura 2.9	Conjuntos Fuzzy resultante da variável linguística gorjeta.....	25
Figura 2.10	Componentes de uma árvore de decisão.....	26
Figura 2.11	Exemplo de dados de treinamento.....	28
Figura 2.12	Exemplo de árvore de decisão.....	29
Figura 2.13	Busca de estados do problema do presente.....	30
Figura 2.14	Construção da árvore em nível parcial.....	34
Figura 2.15	Árvore de decisão final.....	35
Figura 2.16	Arquivo ARFF.....	38
Figura 2.17	WEKA GUI Chooser.....	39
Figura 2.18	Seleção da classe e dos atributos preditivos.....	40
Figura 2.19	Seleção do algoritmo de mineração de dados.....	41
Figura 2.20	Árvore de decisão minerada pelo WEKA.....	42
Figura 3.1	Localização dos municípios de Belém e Ananindeua.....	43
Figura 3.2	Amostra dos dados levantados pela FEC/ITEC/UFPA.....	44
Figura 3.3	Variável linguística de entrada – Área.....	46
Figura 3.4	Variável linguística de entrada – Padrão.....	46
Figura 3.5	Variável linguística de entrada – Idade.....	47
Figura 3.6	Variável linguística de entrada – Local.....	48
Figura 3.7	Variável linguística de entrada – Cidade.....	49
Figura 3.8	Variável linguística de entrada – Vagas.....	50
Figura 3.9	Variável linguística de saída – Valor_m ²	50

Figura 3.10	Resultado da leitura do conjunto de dados pelo software WEKA.	53
Figura 3.11	Resumo da criação da árvore de decisão pelo software WEKA....	54
Figura 3.12	Árvore de decisão criada pelo software WEKA.....	55
Figura 3.13	Regras Fuzzy criadas a partir da árvore de decisão.....	56
Figura 4.1	Valores calculados e observados no mercado do grupo 1.....	59
Figura 4.2	Valores calculados e observados no mercado do grupo 2.....	61
Figura C.1	Árvore de decisão gerada.....	72

LISTA DE TABELAS

Tabela 2.1	Tabela de decisão para o problema do presente.....	30
Tabela 2.2	Condições para a realização do jogo.....	33
Tabela 2.3	Base de dados.....	39
Tabela 3.1	Variáveis linguísticas e suas funções de pertinência.....	45
Tabela 3.2	Conjunto de treino da árvore de decisão.....	53
Tabela 4.1	Primeiro grupo de dados submetido ao modelo Fuzzy.....	57
Tabela 4.2	Valores calculados e observados do grupo 1.....	58
Tabela 4.3	Segundo grupo de dados submetido ao modelo Fuzzy.....	60
Tabela 4.4	Valores calculados e observados do grupo 2.....	60
Tabela A.1	Dados utilizados para a geração da árvore de decisão.....	69
Tabela B.1	Resultados obtidos pelo modelo proposto.....	71
Tabela I.1	Dados do levantamento de campo.....	73
Tabela II.1	Dados inseridos no modelo.....	77

NOMENCLATURA

ABNT	ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS
ARFF	FORMATO DE ARQUIVO DE RELAÇÃO DE ATRIBUTOS
GNU	LICENÇA PÚBLICA DE SOFTWARE LIVRE
INMETRO	INSTITUTO NACIONAL DE METROLOGIA
NBR	NORMA BRASILEIRA
WEKA	AMBIENTE WAIKATO PARA ANÁLISE DO CONHECIMENTO

CAPÍTULO 1

INTRODUÇÃO

1.1 – MOTIVAÇÃO

O processo de avaliação de imóveis envolve diversos parâmetros referentes às variáveis selecionadas para representar o comportamento do mercado imobiliário. Estas variáveis, denominadas independentes ou de entrada, se relacionam com o valor do imóvel, que é a variável dependente ou de saída, geralmente de forma não-linear. Em trabalhos publicados pelo meio acadêmico e profissional, verifica-se que alguns modelos lineares têm sido utilizados adequadamente para descrever o comportamento do mercado imobiliário. Contudo, uma grande parte desses processos, independente da sua natureza, apresentam comportamentos não lineares, muitas vezes não contínuos, que levam a modelos complexos, exigindo soluções analíticas e/ou numéricas.

Uma das limitações da regressão linear está concentrada no mapeamento de dados através de uma aproximação linear de dados de mercado. Esta aproximação pode não refletir o real valor de mercado do imóvel avaliado, especialmente quando os dados são altamente dispersos e com características intrínsecas e extrínsecas das mais variadas.

A aplicação de modelos de decisão, na avaliação de imóveis, exige o atendimento aos seus pressupostos básicos para que a análise seja validada. GONZÁLEZ *et al.* (2000) realizou uma análise conceitual das dificuldades encontradas na determinação dos modelos de regressão, encorajando os estudiosos a buscar fundamentos em outras metodologias, como as Redes Neurais Artificiais (RNA) e a Lógica Fuzzy.

Este trabalho objetiva contribuir com os métodos usuais de avaliação de imóveis, aprimorando o arbítrio do avaliador, nas inferências sobre os valores das variáveis, quando as informações são difusas, propondo um modelo Fuzzy de avaliação.

A utilização de árvores de decisão é proposta como uma forma de simplificação da quantidade de regras Fuzzy utilizadas, tornando viável a implementação prática do modelo.

1.2 – OBJETIVOS

1.2.1 – Objetivo Geral

Propor um modelo de avaliação de imóveis, que incorpore as imprecisões e subjetividades inerentes às variáveis que compõem o mercado imobiliário, para utilização por profissionais que atuam na área.

1.2.2 – Objetivos Específicos

- Levantar as necessidades do mercado de avaliação imobiliária;
- Criar o modelo Fuzzy para avaliação de imóveis;
- Simplificar o número de regras estabelecidas pelo modelo, utilizando Árvore de Decisão;
- Submeter ao modelo criado, dados reais do mercado imobiliário;
- Validar o modelo criado comparando seus resultados aos obtidos diretamente no mercado imobiliário.

1.3 – CONTRIBUIÇÕES DA DISSERTAÇÃO

A NBR 14653-2, da ABNT – Associação Brasileira de Normas Técnicas, prevê no item 8.2.1.4.3 – Tratamento Científico, a utilização de outras ferramentas, além da regressão linear, para a aferição do comportamento do mercado, citando, dentre as metodologias para o tratamento dos dados, a utilização de redes neurais artificiais, a análise envoltória de dados e a regressão espacial.

No entanto, a maioria dos métodos existentes hoje, para mensurar o valor de imóveis, utilizam metodologias estatísticas que são determinísticas ou probabilísticas.

A aplicação da Regressão Linear exige o atendimento aos seus pressupostos básicos para que a análise seja validada. Neste sentido, a utilização desta metodologia encontra dificuldades em questões de grande importância e que têm sido discutidas no meio acadêmico, que são a auto correlação espacial, o desconhecimento da forma funcional para o modelo a ser adotado e a criação de escalas para as variáveis qualitativas. Em GONZÁLEZ *et al.* (2000) é feita uma análise conceitual das dificuldades encontradas

na determinação dos modelos de regressão linear, encorajando os estudiosos a buscar fundamentos em outras metodologias.

Em trabalhos publicados pelo meio acadêmico, verifica-se que alguns modelos lineares têm sido utilizados adequadamente para descrever o comportamento do mercado imobiliário, contudo, uma grande parte destes processos, independente da sua natureza, apresentam comportamentos não lineares e que levam a modelos complexos exigindo soluções analíticas e/ou numéricas. Quando se utiliza a transformação de variáveis buscando a linearização dos valores, e principalmente quando esta prática é aplicada à variável dependente, essa complexidade restringe e dificulta o conhecimento e a análise dos próprios processos.

Uma das limitações da regressão linear está concentrada no mapeamento de dados através de uma aproximação linear de dados de mercado. Esta aproximação pode não refletir o valor de mercado do imóvel avaliado, especialmente quando os dados são altamente dispersos e com características intrínsecas e extrínsecas das mais variadas. Neste sentido, é clara a necessidade de novas técnicas visando representar os processos de avaliação de imóveis (BOLINELLI *et al.*, 2015).

Quando há a necessidade de considerar aspectos subjetivos de difícil mensuração, no valor de um imóvel, há a dificuldade de se trabalhar com aproximação linear.

Na Avaliação Imobiliária há a necessidade de se considerar aspectos subjetivos, ambíguos, imprecisos, de difícil mensuração para determinar o valor de um imóvel.

Dentro deste cenário de imprecisão, abstrato, utilizando variáveis difusas é que convive o segmento do mercado de avaliação imobiliária, necessitando trabalhar com toda essa subjetividade e onde a utilização de novas ferramentas, como a Lógica Fuzzy, se apresenta mais adequada (DUARTE *et al.*, 2006).

Várias abordagens já foram apresentadas na linha de avaliação de imóveis utilizando lógica Fuzzy, das quais destacam-se, pelo seu pioneirismo ou pela visão prática do processo, as seguintes:

DUARTE *et al.* (2006) foram um dos precursores na proposta de uma nova forma de modelar o mercado imobiliário, levando os conceitos científicos já utilizados na Engenharia de Avaliações para a Lógica Fuzzy. Realizaram a modelagem de apartamentos, utilizando cinco atributos como variáveis de entrada, gerando o valor global do imóvel. Por ser um estudo introdutório, das 405 regras Fuzzy possíveis de serem criadas, foram selecionadas apenas 61 para serem utilizadas no modelo.

MELLO (2012) apresentou um método para avaliação de imóveis com uso da Lógica Fuzzy e comparou os resultados com os obtidos através do Custo de Reprodução, calculados através de uma Árvore de Avaliação Econômica. Realizou a modelagem de 1 apartamento e 6 glebas de terra, utilizando dois atributos como variáveis de entrada, gerando o valor global do imóvel. Utilizou todas as 25 regras Fuzzy estabelecidas. Na comparação dos resultados obteve diferenças entre 0,81% e 6,98%.

BOLINELLI *et al.* (2015) apresentaram um método que utiliza a Lógica Fuzzy como ferramenta para estimar a influência de variáveis ambientais nos valores dos imóveis urbanos. Realizaram a modelagem de 25 unidades de um empreendimento residencial, utilizando três atributos como variáveis de entrada, gerando o valor global do imóvel. Utilizaram todas as 27 regras estabelecidas. Os resultados foram comparados com os valores obtidos por Inferência Estatística, calculados através do software SISDEA, obtendo diferenças na ordem de 1,09%.

PELLI *et al.* (2015) realizaram a modelagem Fuzzy de 3 unidades de um empreendimento residencial, utilizando cinco atributos como variáveis de entrada, gerando o valor global do imóvel. Das 3.125 regras estabelecidas, foram selecionadas para uso no modelo apenas 25. Comparou os resultados com os valores obtidos através do Método Comparativo de Dados de Mercado (MCDDM), obtendo diferenças na ordem de 10% a 15%.

Observa-se que, no período dos trabalhos aqui estudados (2006 a 2015) até a data de elaboração desta dissertação, não foram encontradas propostas de métodos de modelagem Fuzzy para Avaliação Imobiliária, que pudessem realmente serem utilizadas, de uma forma prática, pelos profissionais da área.

Um dos fatores observados, que pode estar influenciando nessa dificuldade de implementação é o fato da quantidade de regras Fuzzy, estabelecidas pelo modelo gerado, crescer de forma geométrica, diretamente proporcional ao número de atributos selecionados para avaliação do imóvel, requerendo do especialista na área, que além do trabalho de selecionar os atributos para avaliar os imóveis da sua região, precisa fazer um grande esforço adicional na seleção das regras mais significativas, dentre o total gerado, para serem utilizadas na prática, esforço este que nem sempre é realizado por uma só pessoa. Desta forma temos a seguinte questão norteadora para este trabalho:

Como desenvolver um modelo de avaliação de imóveis, que incorpore as imprecisões e subjetividades, inerentes às variáveis que compõem o mercado imobiliário, para utilização por profissionais que atuam na área?

1.4 - ORGANIZAÇÃO DO TRABALHO

No capítulo 1 buscou-se enfatizar as motivações e objetivos que levaram à proposta de um modelo Fuzzy para avaliação de imóveis, para posterior utilização nas atividades diárias dos profissionais que atuam na área.

No capítulo 2 apresenta-se uma revisão da literatura acerca dos métodos de avaliação imobiliária praticados no Brasil. É feita, também, uma revisão teórica sobre a Lógica Fuzzy. A Lógica Fuzzy tem como principal objetivo a modelagem computacional do raciocínio humano, impreciso, subjetivo, ambíguo e vago, refletindo a maneira como as pessoas pensam. Depois é feita uma revisão teórica sobre classificação de dados utilizando Árvore de Decisão. Os classificadores baseados na árvore de decisão são um dos ramos na área de inteligência artificial, eles pertencem ao campo de aprendizagem de máquina devido sua capacidade de aprender, através de exemplos, com o objetivo de classificar registros em uma base de dados. Esta técnica de classificação é utilizada na base de dados das regras Fuzzy para a seleção das regras mais significativas que serão utilizadas no modelo proposto. Neste capítulo é feita também a apresentação do software WEKA, utilizado para a geração da Árvore de Decisão.

A metodologia empregada é apresentada no capítulo 3, mostrando-se os passos para a criação do modelo Fuzzy proposto, os dados que foram tomados como base nesta criação, a seleção das regras Fuzzy mais significativas através de uma árvore de decisão, utilizando o software WEKA e a implementação final do modelo, utilizando o ambiente de simulação de processos Matlab.

No capítulo 4 são mostrados os resultados obtidos, onde são inferidos ao modelo criado, dados reais do mercado imobiliário, obtendo-se o valor do metro quadrado desses imóveis. Para validação do modelo proposto é feita uma comparação dos resultados calculados com aqueles obtidos diretamente no mercado imobiliário.

No capítulo 5 apresentam-se as principais conclusões e sugestões para a continuação do trabalho em etapas posteriores.

No apêndice A mostra-se o conjunto de 75 dados do Anexo A, adequado para ser utilizado como conjunto de treino/teste para a geração da Árvore de Decisão, pelo software WEKA. Nestes dados constam os atributos que foram selecionados como entradas e a classe de saída, valor do metro quadrado, com os seus valores transformados em funções de pertinência.

No apêndice B mostram-se as planilhas dos resultados obtidos pelo modelo Fuzzy proposto, na avaliação de dois grupos de 12 imóveis cada, que estão no Anexo A.

No Apêndice C mostra-se a Árvore de Decisão gerada pelo software WEKA, para a definição das novas regras utilizadas pelo modelo proposto por este trabalho.

No anexo A estão os dados constantes de um levantamento de campo realizado pela Faculdade de Engenharia Civil/ITEC/UFPA, totalizando 75 apartamentos, nos municípios de Ananindeua e Belém, cujos dados foram utilizados para o treinamento da árvore de decisão, utilizada neste trabalho.

No anexo B estão os dados constantes do mesmo levantamento de campo realizado pela Faculdade de Engenharia Civil/ITEC/UFPA, totalizando 24 apartamentos, cujos dados foram inferidos no modelo proposto, para a obtenção do valor do metro quadrado desses imóveis. Estes dados não fizeram parte do conjunto de dados usados para treino e teste da árvore de decisão, mostrados no Anexo A.

CAPÍTULO 2

REVISÃO DA LITERATURA

2.1 – AVALIAÇÃO IMOBILIÁRIA

A avaliação técnica de imóveis utiliza métodos que visam determinar o valor de mercado de um bem (imóvel). A aplicação do melhor método para se chegar ao valor do imóvel depende das condições que o mercado imobiliário oferece ao avaliador. O fato que deve ser observado, neste momento, são as informações que se têm disponíveis e a finalidade das avaliações. Assim, pode-se dizer que a escolha do método para obter o valor de um imóvel é dependente das informações disponíveis e do nível de rigor que se deseja ter (ABNT, 2011).

A primeira norma brasileira de avaliação de imóveis surgiu em 1952 e foi elaborada pelo departamento de engenharia da Caixa Econômica Federal. A primeira Norma Brasileira para Avaliação de Imóveis Urbanos da Associação Brasileira de Normas Técnicas (ABNT) surgiu em 1977 sob a sigla NB502. Em 1980 esta Norma sofreu sua primeira revisão e foi registrada no INMETRO como NBR5676. Um novo projeto de revisão da NBR5676 foi elaborado por uma comissão de avaliadores. Com este projeto de revisão a NBR5676 (1989) foi transformada na NBR14653-2, a parte II de um conjunto de Normas de avaliação de bens (MANUAL DE PERÍCIAS, 2017).

Atualmente, a avaliação de imóveis urbanos no Brasil é normatizada pela NBR 14653-2 (ABNT, 2011), publicada em maio de 2004 e atualizada em 2011.

A NBR 14653-2 (ABNT, 2011) tem o objetivo de fixar as diretrizes para a avaliação de imóveis urbanos, quanto à classificação de sua natureza, descrição das atividades básicas, do uso das metodologias, das especificações da avaliação, e suporte para laudos e pareceres técnicos de avaliação. Além disso, a norma apresenta ainda, em seu texto, a descrição de vários métodos possíveis que podem ser aplicados nos processos de avaliação.

2.2 - MÉTODOS DE AVALIAÇÃO DE IMÓVEIS

A seguir serão abordados os métodos para avaliação de imóveis apresentados na atual Norma Brasileira de Imóveis Urbanos – NBR14653-2 e na literatura técnico-científica.

2.2.1 – Método comparativo de dados de mercado

O método comparativo de dados de mercado é aquele que define o valor do imóvel através da comparação deste com outros imóveis assemelhados quanto às suas características. Para tanto, é fundamental a existência de imóveis para comparação, ou seja, uma amostra de dados do mercado imobiliário formada pelos chamados imóveis de referência. Ao utilizar este método, deve-se fazer o tratamento estatístico dos dados pesquisados. São duas as modalidades de tratamento estatístico das amostras pesquisadas: a Metodologia Científica, que emprega a Estatística Inferencial e vem sendo mais utilizada por sua maior confiabilidade e a Homogeneização de Valores, que emprega a Estatística Descritiva.

A amostragem é formada pelos imóveis de referência, que são imóveis semelhantes ao imóvel avaliado para os quais se conhece o valor e é na maioria das vezes, tomada nas regiões circunvizinhas ao imóvel avaliado. Isso ocorre pela necessidade de se incluir no processo de avaliação a localização do imóvel, uma característica fundamental para o cálculo de seu valor. Logo, para a utilização do método comparativo dos dados de mercado, o avaliador deve ter conhecimento dos dados de mercado da região que o imóvel avaliado se encontra, o que pode vir a ser um problema para investidores e avaliadores provenientes de outras regiões.

De todos os métodos existentes, o método comparativo de dados de mercado é o mais utilizado para a avaliação de imóveis no Brasil, sejam eles residenciais ou comerciais. Apesar disso, há casos para os quais não existem imóveis de referência que possibilitem o emprego deste método como, por exemplo, o único *shopping Center* em uma cidade de médio porte ou um único galpão de uma cidade de pequeno porte. Nestes casos, deve-se optar por outros métodos de avaliação, como os métodos econômico-financeiros.

2.2.2 – Método comparativo do custo de reprodução

Neste método, o valor de benfeitorias é apropriado através da reprodução dos custos de seus componentes, com base em um orçamento detalhado, devendo-se justificar e quantificar os efeitos do desgaste físico e do obsolescimento funcional das benfeitorias. Deve-se incluir ainda o valor do terreno e, quando for o caso, o fator de comercialização.

Segundo a NBR14653-2, entende-se por fator de comercialização, ou coeficiente de valor adicionado, a razão entre o valor de mercado e o custo de reedição de um imóvel assemelhado. Esse fator considera a vantagem financeira que uma benfeitoria já pronta tem sobre uma a ser construída. Porém, dependendo da conjuntura do mercado na época da avaliação, este pode ser maior ou menor que a unidade.

O método comparativo do custo de reprodução por si só, nada tem a ver com a rentabilidade e a viabilidade do investimento. Mas, ao se determinar o fator de comercialização, já se está, de certo modo, avaliando grosseiramente os riscos do investimento, ou pelo menos se os ativos daquela região tendem a valorizar, quando o fator de comercialização for maior que um, ou a desvalorizar, quando o fator de comercialização for menor que um.

A principal desvantagem deste método é o fato de que neste determina-se o custo de reprodução das benfeitorias e não seu valor de mercado. O fator de comercialização busca minimizar este problema aproximando o custo de reprodução das benfeitorias do valor de mercado, mas não é simples de ser obtido.

2.2.3 – Método da renda

O método da renda é aquele que apropria o valor do imóvel e de suas partes constitutivas, com base na capitalização presente da sua renda líquida, seja ela real ou prevista. Para a sua utilização, precisam ser determinados o período de capitalização e a taxa de desconto a ser aplicada nos fluxos de caixa, que devem ser devidamente fundamentados e expressamente justificados pelo avaliador, como exige a Norma. Este método é fundamentado no princípio de que o valor de uma determinada propriedade é uma função de sua capacidade de gerar renda, o que geralmente é o aluguel.

Segundo MENDONÇA *et al.* (1998), devido à sua flexibilidade que permite a grande generalização do seu uso, nos casos de perícias judiciais, este método tem sido recomendado pela própria jurisprudência dos tribunais do poder judiciário, que consagra

como justa a taxa de desconto de 12% ao ano para imóveis comerciais. Porém, os trabalhos que vem sendo realizados com provam que não existe uma taxa de desconto pré-definida que indique com segurança a renda mensal esperada para um determinado imóvel, em virtude da constante instabilidade econômica do país. Não há, no Brasil, taxas de desconto fixadas para imóveis residenciais, comprovando-se que, geralmente, quanto maior e mais luxuoso é o imóvel residencial, menor será a taxa interna de retorno.

Ao se utilizar o método da renda, a rentabilidade do imóvel é explícita, o que o torna comparável a investimentos em outros tipos de imóveis e também em outros ativos, como ações, caderneta de poupança, etc. Este método é uma ferramenta adequada para se analisar a viabilidade do investimento.

2.2.4 – Método involutivo

O método involutivo é baseado em modelo de estudo de viabilidade técnico-econômica para apropriação do valor do terreno, alicerçado no seu aproveitamento eficiente, mediante hipotético empreendimento imobiliário compatível com as características do imóvel e com as características de mercado. Esta Norma fala sobre a utilização deste método apenas para a avaliação de terrenos, mas nada impede, além das limitações da Norma, que este seja aplicado a outros bens. No projeto de revisão da Norma, o conceito de método involutivo será alterado, não mais restringindo-o apenas a terrenos, mas validando-o a qualquer bem.

Para a aplicação do método involutivo deve-se considerar a receita provável da comercialização das unidades hipotéticas com base em preços obtidos em pesquisas, todas as despesas inerentes à transformação do terreno, a margem de lucro do empreendedor, as despesas de comercialização, os prazos viáveis ao projeto, à execução e à comercialização, mediante taxas financeiras operacionais reais (NBR14653-2). É importante observar que a receita provável da comercialização das unidades hipotéticas, quando houver, deve ser calculada com base em preços obtidos em pesquisas, o que significa que estes preços podem ser obtidos pelo método comparativo de dados de mercado.

Assim como no método da renda, a melhor maneira de se determinar a taxa de desconto apropriada é fazendo um estudo sobre os riscos do investimento, e comparando sua rentabilidade com a de imóveis similares, sempre que possível. É muito importante que o avaliador conheça o mercado ao qual pertence o imóvel que está sendo avaliado.

Atualmente, o método involutivo é amplamente utilizado na avaliação de glebas urbanizáveis na falta de imóveis de referência. Glebas urbanizáveis são grandes extensões de terreno passíveis de receber obras de infraestrutura urbana (NBR14653-2), podendo ser urbanas ou não urbanas, residenciais, comerciais e/ou industriais ou mistas. Quando não existem outras glebas cujas características se assemelham ao imóvel avaliado, servindo de base de comparação, a utilização do método comparativo de dados de mercado torna-se inviável. Pelo método involutivo, o valor da gleba é definido através do estudo do seu melhor aproveitamento, considerando as características socioeconômicas da região.

Para se obter o valor do imóvel pelo método involutivo, utiliza-se o fluxo de caixa descontado no qual as entradas de caixa são os recebimentos esperados para o empreendimento, considerando-se o melhor aproveitamento do bem, ou seja, o aproveitamento eficiente. As saídas de caixa são todos os desembolsos de execução, transformação e comercialização, incluindo impostos. Este método baseia-se principalmente nos recebimentos que podem vir a ser obtidos pelo imóvel.

2.2.5 – Método residual

O método residual é aquele que define o valor do terreno por diferença entre o valor total do imóvel (terreno e suas benfeitorias) e o valor das benfeitorias. Deve-se também, quando for o caso, considerar o fator de comercialização (NBR14653-2).

2.2.6 – Modelo por regressão

Nas atividades de avaliação imobiliária, como os imóveis são heterogêneos e há diversas características importantes a serem consideradas simultaneamente, é necessário utilizar uma técnica para realizar o ajustamento das diferenças entre os imóveis. Uma das técnicas mais utilizadas atualmente pelos avaliadores e recomendada pela NBR 14.653-2 (GONZÁLEZ, 2002).

Modelos por regressão são modelos matemáticos que relacionam o comportamento de uma variável Y com outra X. A variável X é a variável independente da equação, enquanto Y é a variável dependente das variações de X. Este comportamento é representado por uma reta chamada de reta de regressão. Esta reta explica de forma geral e teórica a relação entre X e Y. Isto significa que os valores observados de X e Y

nem sempre serão iguais aos valores de X_1 e Y_1 estimados pela reta de regressão. Haverá sempre alguma diferença.

Esta diferença, em estatística, é chamada de erro ou desvio. O processo de regressão significa que os pontos plotados no gráfico são definidos, modelados ou regredidos, a uma reta, que corresponde à menor distância possível entre cada ponto plotado e a reta, conforme mostrado na Figura 2.1. Em outras palavras, busca-se reduzir ao mínimo possível os somatórios dos desvios entre Y e Y_1 . Este modelo é conhecido também como Inferência Estatística ou Modelagem Estatística.

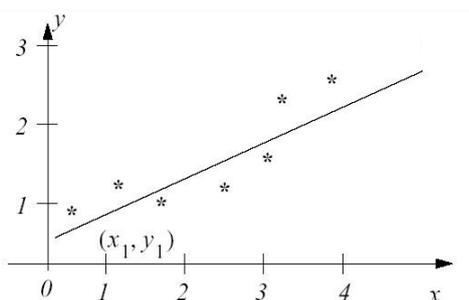


Figura 2.1 – Reta de regressão.
Fonte: DUARTE (2006).

Na Avaliação de Imóveis, a regressão busca verificar a relação entre a variável dependente (o valor do imóvel) e as independentes, tais como: área do terreno, área construída, idade do bem, localização, entre outras, de tal forma que possa ser traçada uma curva ou reta, representada matematicamente por uma equação (modelo), que acomode ou ajuste a nuvem de pontos formada pelo conjunto de dados. A curva ou reta que ajusta os dados amostrais, possibilitando estimar o valor da variável dependente (Y), correspondente a uma ou mais variáveis independentes (X_n) conhecidas, é denominada **Regressão de Y para X**. As respostas ou saídas dos modelos de regressão apresentam-se contínuas (DUARTE, 1999). Este tipo de modelagem é o mais utilizado no mercado de Avaliação de Imóveis e segundo MANUAL DE PERÍCIAS (2015). Para as avaliações de imóveis, a Caixa Econômica Federal - CEF, exige que sejam feitas através de inferência estatística.

2.3 – LÓGICA FUZZY

O que quer dizer o termo ‘Fuzzy’?

A palavra ‘Fuzzy’ nos dicionários de língua Inglesa significa borrado, indistinto (‘blurred’); imprecisamente definido, confuso, vago. (No Brasil, alguns chamam de ‘nebuloso’, ‘difuso’).

Algumas razões para usar lógica Fuzzy (WANG, 1997):

- Ela é conceitualmente fácil de entender;
- Ela é flexível;
- Ela pode modelar funções não lineares de arbitrária complexidade;
- Sistemas Fuzzy podem ser construídos baseando-se somente na experiência de especialistas;
- Lógica Fuzzy pode ser misturada com técnicas de controle convencional;
- Lógica Fuzzy é baseada em linguagem natural.

Tomando como exemplo a frase a seguir: “Embora o transformador esteja um pouco carregado, pode-se utilizá-lo por um tempo”.

Tem-se que nesta frase existem dois termos (um pouco e um tempo) bastantes subjetivos e difíceis de representar. Para um especialista em eletricidade, que esteja conversando com outro especialista, o entendimento seria normal; no entanto, durante o processo de modelagem matemática do problema, fica bastante complicado representar esse tipo de informação. Dessa forma, uma maneira de tentar solucionar o processo de representação de conhecimento impreciso é através da Lógica Fuzzy. A Lógica Fuzzy tem como principal objetivo a modelagem computacional do raciocínio humano, impreciso, ambíguo e vago (MARRO, 2010).

A Lógica Fuzzy (também chamada de lógica multivalorada) foi primeiramente introduzida em 1930 pelo filósofo e lógico polonês Jan Lukasiewicz. Através do estudo de termos do tipo alto e baixo, velho e quente, ele propôs a utilização de um intervalo de valores $[0,1]$ que indicaria a possibilidade de uma declaração ser verdadeira ou falsa. Em 1937, o filósofo Max Black propôs a ideia de que continuidade descrevia graus. Ele definiu o primeiro conjunto Fuzzy e descreveu algumas ideias básicas de operações com conjuntos Fuzzy.

Em 1965, Lofti Zadeh publicou o artigo Fuzzy Sets, que ficou conhecido como a origem da Lógica Fuzzy. Na realidade, Zadeh redescobriu a ideia de fuzzyficação,

identificou e explorou tal conceito, assim como lutou por ele. Portanto, Zadeh ficou e ainda é conhecido como o “mestre” da Lógica Fuzzy.

A característica especial da Lógica Fuzzy é a de representar uma forma inovadora no manuseio de informações imprecisas, de forma muito distinta da teoria de probabilidades. A Lógica Fuzzy provê um método de traduzir expressões verbais, vagas, imprecisas e qualitativas, comuns na comunicação humana em valores numéricos (MARÇAL e SUSIN, 2005).

Deste modo, é possível se converter a experiência humana em uma forma compreensível pelo computador. Assim, a tecnologia, utilizando o enfoque Fuzzy, tem um imenso valor prático, pela qual se torna possível a inclusão da experiência de operadores humanos, os quais controlam os processos e plantas industriais, em controladores computadorizados, possibilitando estratégias de tomadas de decisão em problemas complexos ou que exijam tempo de solução reduzido. O sucesso atualmente reconhecido de sistemas de modelagem e controle baseados em Lógica Fuzzy tem comprovado sua utilização como mais uma ferramenta (ou tecnologia) para a resolução de problemas de engenharia de controle industrial, manufatura, manutenção, comunicação homem-máquina e em sistemas de tomadas de decisão.

2.4 - FUNDAMENTOS DA LÓGICA FUZZY

A teoria clássica de conjuntos permite o tratamento de classes de objetos e suas inter-relações em um universo definido. Nessa teoria, a *pertinência* de um dado elemento com relação a um conjunto refere-se ao fato de tal elemento pertencer ou não a esse conjunto (FEITOSA, 1992).

De forma ilustrativa, considere o gráfico apresentado no lado esquerdo da Figura 2.2, que representa um exemplo típico da teoria clássica e descreve a altura de uma pessoa, em centímetros, através de três conjuntos: *baixo*, *médio* e *alto*. Nesse exemplo, dado um elemento x qualquer, o mesmo pertencerá a um dos conjuntos do gráfico; por exemplo, se $x = 169$, então x pertence ao conjunto *baixo* e não aos demais, ou seja, um elemento pertence ou não a um determinado conjunto e, além disso, tal elemento não pertence a mais de um conjunto.

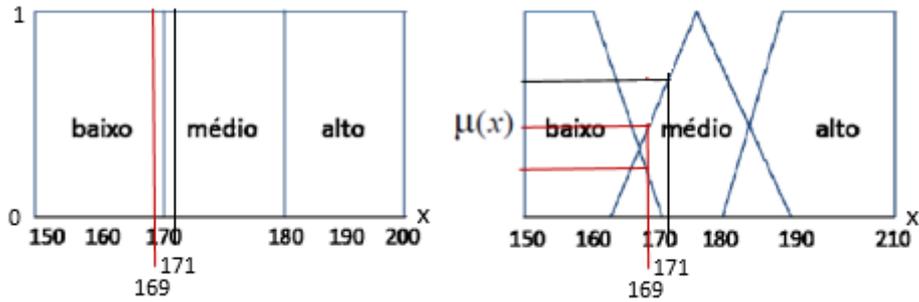


Figura 2.2 - Representação na forma de conjuntos da altura de uma pessoa, sob o ponto de vista da Lógica convencional (à esquerda) e do da Lógica Fuzzy (à direita).

Fonte: FEITOSA (1992).

Ao contrário da Lógica convencional, a Lógica Fuzzy utiliza a idéia de que todas as propriedades mensuráveis (temperatura, altura, velocidade, etc.) admitem *graus de pertinências*. Com isso, a Lógica Fuzzy tenta modelar o senso das palavras, tomada de decisão ou o senso comum do ser humano.

Ainda tomando como exemplo o gráfico da Figura 2.2, dados dois elementos $x_1 = 169$ e $x_2 = 171$, se a Lógica clássica for utilizada, esses dois elementos pertencem a classes diferentes, x_1 pertencendo à classe baixo e x_2 à classe média (lógica convencional). No entanto, na realidade fica difícil de dizer que uma pessoa com 169 cm e outra com 171 cm pertencem a classes diferentes.

Por outro lado, na Lógica Fuzzy, tanto o x_1 quanto o x_2 têm graus de pertinências parciais aos dois conjuntos *baixo e médio*, que podem variar de 0 a 1. Em outras palavras, enquanto que a tomada de decisão na Lógica convencional seria como a da Eq. (2.1), onde a saída assume somente dois valores zero e um, a da Lógica Fuzzy seria como a da Eq. (2.2), considerando um conjunto A e um elemento x com relação a esse conjunto, onde a saída assume vários valores de acordo com a função de pertinência $\mu(x)$.

$$f(x) = \begin{cases} 1 & \text{se, e somente se, } x \in A \\ 0 & \text{se, e somente se, } x \notin A \end{cases} \quad (2.1)$$

$$\mu(x) = \begin{cases} 1 & \text{se, e somente se, } x \in A \\ 0 & \text{se, e somente se, } x \notin A \\ 0 \leq \mu(x) \leq 1 & \text{se } x \text{ pertence parcialmente a } A \end{cases} \quad (2.2)$$

Dessa forma, a Lógica Fuzzy pode ser considerada como um conjunto de princípios matemáticos para a representação do conhecimento baseado no grau de

pertinência dos termos ($\mu(x)$). Como pode ser observado na Eq. (2.2), o intervalo de pertinência é $[0,1]$, onde 0 significa que um elemento não pertence a um determinado conjunto, 1 significa completa pertinência ao conjunto, e valores entre 0 e 1 representam graus parciais de pertinências. Assim, na Lógica Fuzzy, um elemento pertence a um conjunto com um certo grau de pertinência, fazendo com que uma determinada sentença possa ser *parcialmente verdadeira* e *parcialmente falsa*. Além do mais, um mesmo elemento pode ter graus de pertinências diferentes de 0 para mais de um conjunto Fuzzy (MARRO, 2010).

Na Lógica Fuzzy a altura de uma pessoa pode ser descrita através de conjuntos Fuzzy, como mostrado no lado direito da Figura 2.2. Onde, $x_1 = 169$ e $x_2 = 171$, percebe-se que ambos têm graus de pertinência 0 para o conjunto *alto* e graus de pertinência parciais para os conjuntos *baixo* e *médio*. É possível notar que, com a Lógica Fuzzy, não existe um limite abrupto que define os elementos que pertencem ou não a um determinado conjunto, como no caso da Lógica Convencional, os graus de pertinência dos elementos possuem variações suaves no intervalo real $[0,1]$, representando, assim, de forma mais realista, o conhecimento humano.

ZADEH (1987) propôs uma caracterização mais ampla, fazendo com que a passagem da pertinência para não pertinência fosse mais lenta e gradual, na medida em que sugere que alguns elementos são mais pertinentes de um conjunto do que outros. O fator de pertinência pode então assumir qualquer valor entre 0 e 1, sendo que o valor (0) indica uma completa exclusão e o valor (1) representa a completa pertinência ou inclusão. Esta generalização aumenta o poder da função característica, que pode assumir várias formas conforme mostrado na Figura 2.3. Por exemplo, para expressar a ideia de que uma temperatura tem seu valor perto de $60\text{ }^\circ\text{C}$ pode-se utilizar uma função de pertinência triangular ilustrada na Figura 2.3a, onde os valores maiores e menores que $60\text{ }^\circ\text{C}$ são distribuídos simetricamente, com máximo grau de pertinência em $60\text{ }^\circ\text{C}$, para sugerir a ideia de que quanto mais perto a temperatura de $60\text{ }^\circ\text{C}$, mais ele se identifica com o valor representado.

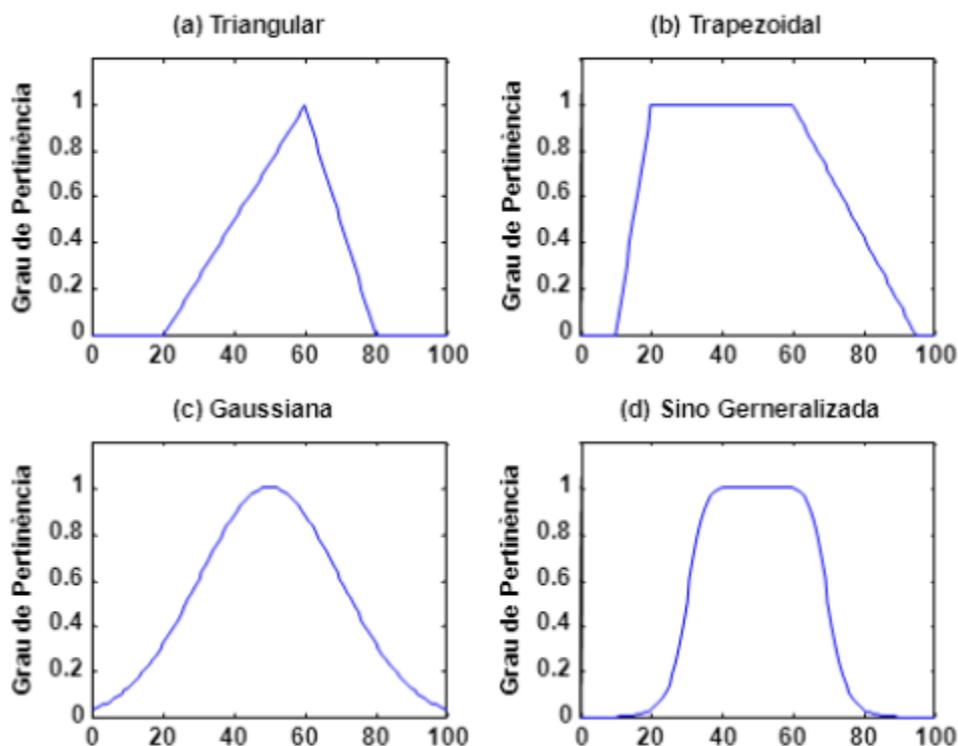


Figura 2.3 – Tipos de função de pertinência.
 Fonte: ZADEH (1987).

Um outro fenômeno pode ser caracterizado pela função trapezoidal ilustrada na Figura 2.3b, a qual expressa o fato de qualquer temperatura entre 20 °C e 60 °C, ter a mesma importância que a temperatura central de 40 °C. Se a necessidade é fazer com que os valores menores que 50 °C tenham pertinência maior dentro do conjunto e, restringir este fato aos valores maiores, pode ser usada a função gaussiana ilustrada na Figura 2.3c, a qual vem de encontro ao que se deseja representar.

2.4.1 - Conjuntos Fuzzy

Como discutido anteriormente, na Lógica clássica os conjuntos são bem definidos, de modo que um elemento pertence ou não a um conjunto; se pertencer, pertence somente a um. Isso evita que ambiguidades apareçam e tornam a lógica mais simples. Ainda considerando o exemplo da utilização de conjuntos para separar pessoas pela altura, uma pessoa com 1,69 m seria considerada uma pessoa de altura mediana, se assim fosse definido, estando apenas nesse conjunto e em nenhum outro; já uma pessoa com 1,71 m faria parte do conjunto das pessoas altas, e somente deste. Todavia, na realidade, fica bem difícil ver que pessoas com uma diferença de altura tão mínima pertencem a conjuntos

diferentes. Por outro lado, pela ótica da Lógica Fuzzy, ter-se-ia as duas pessoas com certo grau de pertinência aos dois conjuntos, variando entre 0 e 1, ou seja, teríamos a tomada de decisão baseada em fatores mais humanos, mais maleáveis. Assim, pode-se concluir que os conjuntos Fuzzy que classificam os elementos de um dado universo são menos rígidos do que aqueles utilizados na teoria clássica visto que eles admitem graus parciais de pertinência.

2.4.2 - Variáveis linguísticas e modificadores

Pode-se considerar uma variável linguística (ou Fuzzy) como uma entidade utilizada para representar de modo linguístico – e, portanto, impreciso – um conceito ou uma variável de um dado problema. Uma variável linguística, diferentemente de uma variável numérica, admite apenas valores definidos na linguagem Fuzzy que está utilizando-se dela. Por exemplo: João é alto.

A variável João está recebendo o valor alto, que é um dos conjuntos Fuzzy definidos para esta variável.

Os modificadores são termos ou operações que modificam a forma dos conjuntos Fuzzy (ou seja, a intensidade dos valores Fuzzy), podendo-se citar, por exemplo, os advérbios muito, pouco, extremamente, quase, mais ou menos, entre outros. Estes podem ser classificados em aumentadores, quando aumentam a área de pertinência de um conjunto Fuzzy, ou, analogamente, diminuidores, quando diminuem a área de pertinência de um conjunto Fuzzy.

2.4.3 - Base de regras

A base de regras tem por objetivo representar de forma sistemática a maneira como o controlador gerenciará o sistema sob sua supervisão.

A forma mais geral de uma regra linguística é:

- **SE** premissa **ENTÃO** consequência.

As premissas também chamadas de antecedentes são associadas com as entradas do controlador Fuzzy e formam a parte das regras representada à esquerda, enquanto as consequências, que também são conhecidas como ações, estão associadas às saídas dos controladores.

Relacionadas às variáveis linguísticas do processo, pode-se estabelecer os antecedentes, ou estados do processo e associá-los com ações de controle. Para exemplificar a utilização da base de regras, toma-se como exemplo o controle de uma caldeira, onde se estabelece uma relação entre a temperatura interna da unidade, com a vazão necessária de água para o seu resfriamento, o que produz as seguintes regras:

- **SE** temperatura é alta **ENTÃO** vazão de água é alta;
- **SE** temperatura é baixa **ENTÃO** vazão de água é baixa;
- **SE** temperatura é estável **ENTÃO** vazão de água é normal.

Aparentemente simples, o exemplo dado é capaz de manter estável a temperatura em uma caldeira, desde que os conjuntos Fuzzy envolvidos tenham uma definição condizente com a realidade. O que vale observar é que para cada estado do processo é relacionada uma ação de controle (VIEIRA, 1999).

2.4.4 – Inferência Fuzzy

A inferência Fuzzy é um processo de avaliação de entradas com o objetivo de, através das regras previamente definidas e das entradas, obter conclusões utilizando-se a teoria de conjuntos Fuzzy. Esse processo pode ser feito através de modelos de inferência, cuja escolha deve levar em consideração o tipo de problema a ser resolvido, obtendo-se assim um melhor processamento. Os modelos de inferência existentes são o estilo Mamdani e o Estilo Sugeno.

2.4.4.1 - Estilo Mamdani

O estilo de inferência Mamdani foi criado pelo professor Ebrahim Mamdani da Universidade de Londres (Reino Unido) em 1975 no contexto do desenvolvimento de sistemas Fuzzy baseando-se em regras de conjuntos Fuzzy no intuito de representar experiências da vida real. Para a construção desse sistema, foi definido um processo de raciocínio dividido em quatro passos: (1) fuzzyficação, (2) avaliação das regras Fuzzy (3) agregação das regras Fuzzy através da máquina de inferência Fuzzy e (4) defuzzyficação, conforme mostrado na Figura 2.4 (WANG, 1997).

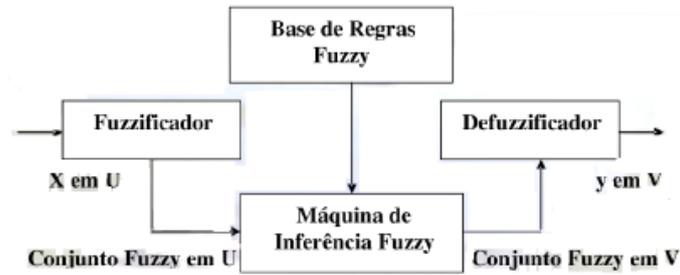


Figura 2.4 – Sistema Fuzzy de Mamdani.
Fonte: WANG (1997).

A Figura 2.4 apresenta os módulos do sistema Fuzzy de Mamdani, onde:

- **Fuzzificador** - converte os valores reais de entrada em um grau de pertinência a conjuntos Fuzzy, para que sejam tratados pela máquina de inferência;
- **Base de regras** - consiste de um conjunto de regras "Se-Então", sendo considerada o “coração” de um sistema Fuzzy, uma vez que todos os outros componentes são usados para implementar as regras de modo eficiente;
- **Máquina de inferência** Fuzzy – combina as regras Fuzzy "SE-Então", existentes na base de regras, no mapeamento de um conjunto Fuzzy de entrada para um conjunto Fuzzy de saída;
- **Defuzzificador** – converte o conjunto Fuzzy, produzido pela máquina de inferência, em um valor real.

2.4.4.2 - Estilo Sugeno

No sistema Fuzzy Takagi-Sugeno, também conhecido como sistema Fuzzy TS, as entradas e saídas são variáveis reais. Entretanto ao invés de considerar as regras Fuzzy SE-ENTÃO no formato de regra do modelo linguístico (sistema Mamdani), estes sistemas usam as regras na seguinte forma:

- Se x é A (premissa), Então $y = B$ (consequente).

Verifica-se que a parte consequente muda de uma descrição que usa termos linguísticos para uma simples fórmula matemática. Esta mudança torna mais fácil combinar as regras. Assim, no sistema Fuzzy TS é obtido uma média ponderada dos valores nas partes ENTÃO das regras (WANG, 1997). A Figura 2.5 mostra a estrutura do sistema Fuzzy Sugeno, onde as entradas e saídas são valores reais, não havendo as etapas de fuzzyficação e defuzzyficação.

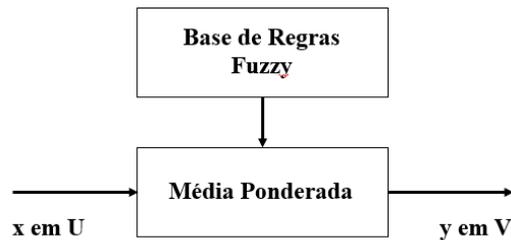


Figura 2.5 – Sistema Fuzzy de Sugeno.
Fonte: WANG (1997).

A Figura 2.5 apresenta os módulos do sistema Fuzzy de Sugeno, onde:

- **Base de regras** - consiste de um conjunto de regras "Se-Então", sendo considerada o “coração” de um sistema Fuzzy, uma vez que todos os outros componentes são usados para implementar as regras de modo eficiente;
- **Média Ponderada** – combina as regras Fuzzy "Se-Então", existentes na base de regras, obtendo uma média ponderada dos valores nas partes Então das regras, no mapeamento do conjunto de entrada real para um conjunto de saída também real.

Neste trabalho utilizaremos a inferência Fuzzy estilo Mandani, que se adequa mais ao intuito de representar experiências da vida real através de termos subjetivos.

2.4.4.3 - Fuzzyficação

Essa etapa obtém o grau de pertinência com que cada entrada pertence a cada conjunto Fuzzy. Cada uma dessas entradas foi previamente limitada no universo de discurso em questão e associada a um grau de pertinência em cada conjunto Fuzzy através do conhecimento do especialista. Então para obter o grau de pertinência de uma determinada entrada *crisp* (não Fuzzy) basta buscar esse valor na base de conhecimento do sistema Fuzzy.

Para ilustração tem-se um exemplo retirado do manual do Fuzzy Logic Toolbox do Matlab, sobre a determinação da *gorjeta* em um restaurante a partir de dados sobre a *qualidade_da_comida* e o *atendimento*. As entradas, qualidade da comida e o atendimento, recebem notas de 0 a 10 e a saída, o valor da gorjeta, varia de R\$ 0,00 a R\$ 30,00.

Os conjuntos Fuzzy e as funções de pertinência para cada uma das variáveis de entrada e saída, são mostrados na Figura 2.6.

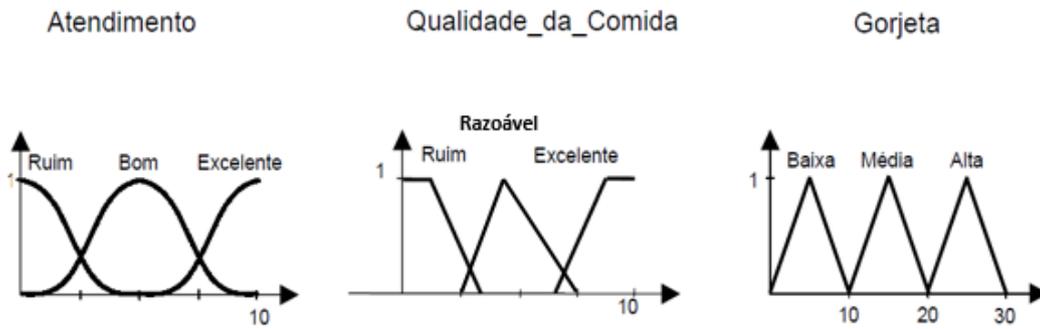


Figura 2.6 – Funções de pertinência definidas para o exemplo da gorjeta.
 Fonte: FUZZY LOGIC TOOLBOX DO MATLAB.

Observa-se na Figura 2.6 que o valor das variáveis linguísticas foram definidas para cada uma das entradas e saída: Atendimento (Ruim, Bom, Excelente), Qualidade da Comida (Ruim, Razoável, Excelente) e Gorjeta (Baixa, Média, Alta).

A Figura 2.7 mostra o grau de pertinência das entradas, para quando uma nota de 7,5 for dada para o Atendimento e uma de 8,0 for dada para a Qualidade da Comida. Os valores para as funções de pertinência do Atendimento são $\mu(7,5) \text{ Excelente} = 0,4$ e $\mu(7,5) \text{ Bom} = 0,3$ e da Qualidade da Comida é $\mu(8,0) \text{ Excelente} = 0,7$.

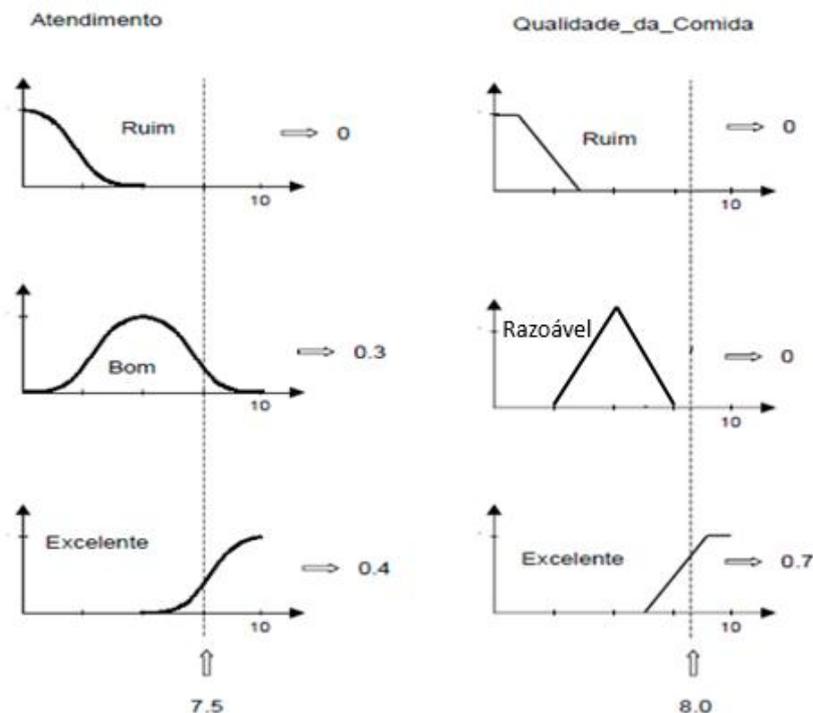


Figura 2.7 – Grau de pertinência das entradas para o exemplo da gorjeta.
 Fonte: FUZZY LOGIC TOOLBOX DO MATLAB.

2.4.5 – Avaliação das regras Fuzzy

Depois de se obter as entradas fuzzyficadas (excelente e bom para o Atendimento e excelente para a Qualidade da Comida) deve-se aplicá-las nos antecedentes (SE) das regras Fuzzy, obtendo assim o valor do conseqüente (ENTÃO) para cada regra. Para um antecedente composto, os operadores *e* e *ou* são utilizados para obter um único resultado; no caso do operador *ou* é utilizada a operação de união (pega-se o maior grau de pertinência), e, no caso do operador *e*, é utilizada a de interseção (pega-se o menor grau de pertinência). Depois de obter um único valor para o antecedente é necessário obter o valor do conseqüente através de um método de correlação dos mesmos. O método mais comum é conhecido como *clipped*, onde o conseqüente é “cortado” para o nível de valor verdade do antecedente da regra avaliada, ou seja, o valor obtido é simplesmente passado para o conseqüente dessa regra.

No exemplo, com base nos graus de pertinência e nas correlações entre as variáveis linguísticas, obteve-se as seguintes regras:

- SE atendimento é ruim ou comida é ruim, ENTÃO gorjeta é baixa;
- SE atendimento é bom, ENTÃO gorjeta é média;
- SE atendimento é excelente ou comida é excelente, ENTÃO gorjeta é alta.

É importante notar que, a regra 1 não é atendida. Na regra 2, o grau de pertinência do atendimento bom é 0,3 então é aplicado o *clipped*: como tem-se apenas um valor na variável linguística de entrada que consta no antecedente (no caso atendimento), então o mesmo é passado para a variável linguística de saída que consta no conseqüente (no caso, gorjeta). Na regra 3 tem-se operador composto *ou*, entre as entradas atendimento e Qualidade da comida, então é passado para a variável linguística de saída que consta no conseqüente (no caso, gorjeta) o maior dentre os graus de Atendimento e Qualidade, no caso 0,7, conforme mostrado na Figura 2.8-A.

2.4.6 – Agregação das regras Fuzzy

Nessa etapa são agregadas todas as funções de pertinência, resultantes dos conseqüentes de cada regra, em um único conjunto Fuzzy. Para o exemplo em questão, considera-se os graus de pertinência, produzidos pela aplicação das regras Fuzzy 2 e 3, na variável linguística de saída gorjeta, mostrados na Figura 2.8-A. A agregação desses

conjuntos resulta no conjunto Fuzzy mostrado na Figura 2.8-B, que corresponde aos valores finais das funções de pertinência da variável linguística de saída gorjeta.

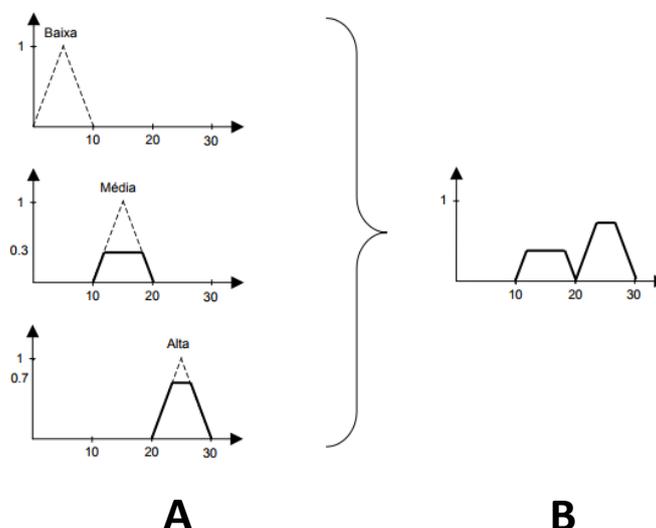


Figura 2.8 - Conjuntos Fuzzy dos graus de pertinência, produzidos pelas regras Fuzzy 2 e 3, na variável linguística de saída gorjeta.

Fonte: FUZZY LOGIC TOOLBOX DO MATLAB.

Na Figura 2.8-B, verifica-se que o conjunto Fuzzy resultante para a variável gorjeta é uma composição dos consequentes de cada regra que foi aplicada.

2.4.7 – Defuzzyficação

O defuzzificador é que pesa as diversas respostas fornecidas pelas regras lógicas e atribui à saída um número. Esse número é a resposta final do sistema Fuzzy e que dirá o que é mais *pertinente* para se fazer.

Os dois métodos mais importantes são: o método baseado no centro de massa, também conhecido como Centróide e o método baseado na média dos máximos das funções de pertinência, também conhecido como Mamdani elaborado por MANDANI. Por um desses métodos encontra-se o valor numérico no eixo “x” mais pertinente.

O método de defuzzyficação mais comum, que foi utilizado no modelo proposto por este trabalho, é o método do Centróide, que obtém o ponto onde uma linha vertical divide ao meio o conjunto agregado, conforme mostrado na Figura 2.9.

A principal vantagem desse método é que o centroide das áreas é encontrado facilmente, quando utilizado com funções de pertinência de pequena complexidade (WANG, 1997).

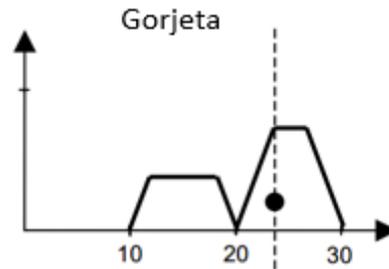


Figura 2.9 - Conjuntos Fuzzy resultante da variável linguística de saída gorjeta, onde é aplicado o método do centroide.

Fonte: WANG (1997).

Este método transforma a saída Fuzzy em saída discreta encontrando a média aritmética entre os centros de gravidade dos conjuntos Fuzzy aos quais o elemento pertence, ponderados pelo grau de pertinência, conforme a Eq. (2.3).

$$Y = \frac{\sum_{i=1}^k \mu_i \cdot Y_i}{\sum_{i=1}^k \mu_i} \quad (2.3)$$

Sendo:

- Y = a saída discreta;
- k = a quantidade de conjuntos Fuzzy a que o elemento pertence;
- μ_i = o grau de pertinência com que o elemento pertence ao i-ésimo conjunto Fuzzy;
- Y_i = o centro de gravidade do i-ésimo conjunto Fuzzy a que o elemento pertence.

2.5 - ÁRVORE DE DECISÃO

As árvores de decisão constroem modelos de classificação ou de regressão na forma de uma estrutura de árvore, transformam um conjunto de dados em subconjuntos cada vez menores, na forma de uma árvore, com o objetivo de converter um grande problema em problemas menores, facilitando assim a tomada de decisão.

Uma árvore é formada por vários componentes: a raiz, que inicializa a árvore, os ramos que são os rótulos com os resultados possíveis de cada teste, os nós internos da árvore, que correspondem a um teste do valor de uma propriedade, as folhas que especificam os valores a serem retornados, se aquela folha for alcançada, como mostra a Figura 2.10.

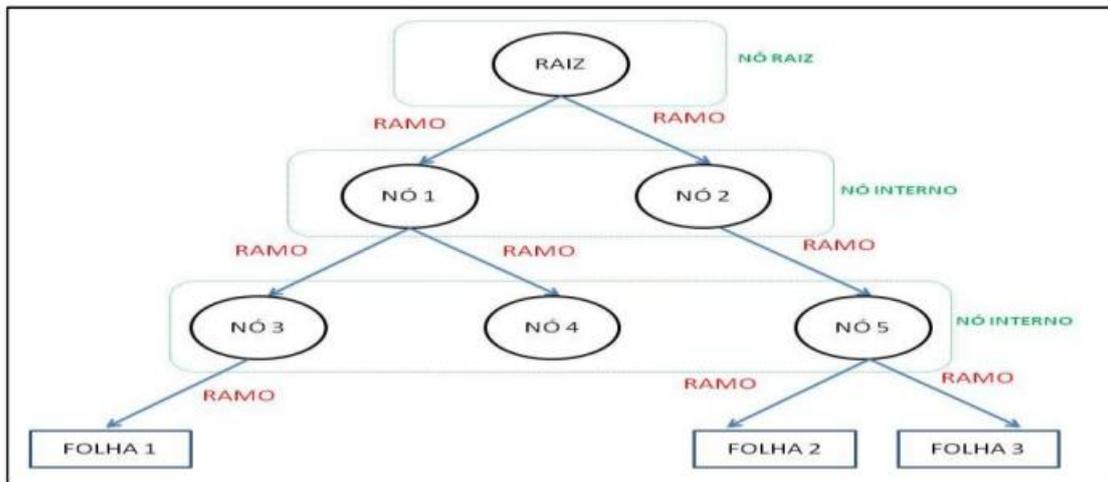


Figura 2.10 – Componentes de uma árvore de decisão.
Fonte: QUINLAN (1986).

De acordo com QUINLAN (1986), as árvores de decisões classificam instâncias, classificando-os da raiz da árvore para baixo até algum nó folha, que prevê a classificação da instância. Cada nó da árvore especifica um teste de algum atributo e cada ramo descendente, a partir desse nó, corresponde a um dos valores possíveis para este atributo. Um exemplo é classificar, iniciando no nó raiz da árvore, testando o atributo especificado por este nó, em seguida, movendo-se para baixo, para o galho da árvore correspondente ao valor do atributo no exemplo dado. Este processo é repetido para a sub-árvore com raiz no novo nó.

A Figura 2.10 é um exemplo de uma árvore de decisão binária. Uma árvore de decisão consiste de nodos e ramos. Cada nodo representa um simples teste ou decisão. No caso de uma árvore binária, a decisão pode ser verdadeira ou falsa. O nodo inicial é comumente referido como nodo raiz. Dependendo do resultado do teste, a árvore poderá se ramificar à esquerda ou à direita em direção a outro nodo. Por fim, o nodo terminal, conhecido também como nodo folha, representado na forma de quadrados, é alcançado, e uma decisão é realizada a uma classe designada. Na construção de árvores de decisão é

normalmente utilizada a convenção de realizar decisões verdadeiras em ramos da direita e decisões falsas no ramo da esquerda.

Para qualquer árvore, todos os caminhos conduzem a um nodo terminal, correspondendo a uma regra de decisão, a qual é uma conjunção de vários testes. Se há múltiplos caminhos para uma dada classe, então os caminhos representam disjunções. Todos os caminhos numa árvore de decisão são mutuamente exclusivos. Para cada novo caso, um e somente um caminho na árvore deverá ser satisfeito.

Segundo QUINLAN (1986), Árvores de Decisão visualizam o domínio da tarefa como uma classificação. A estrutura subjacente consiste em uma coleção de atributos ou propriedades que são utilizadas para descrever casos individuais, cada caso pertencente a um exato conjunto de classes. Os atributos podem ser contínuos ou discretos. O valor de um caso de um atributo contínuo é sempre um número real, enquanto seu valor de um atributo discreto é um pequeno conjunto de valores possíveis para esse atributo.

Árvores de Decisão são modelos estatísticos que utilizam um treinamento supervisionado para a classificação e previsão de dados. Em outras palavras, em sua construção é utilizado um conjunto de treinamento formado por entradas e saídas (SILVA, 2005).

Uma árvore de decisão recebe, como entrada, um objeto ou situação descrita por um conjunto de propriedades e produz uma “decisão” como, por exemplo, “sim” ou “não”. Cada nó interno na árvore corresponde a um teste do valor de uma das propriedades e os ramos do nó são rotulados com os possíveis valores do teste. Cada nó folha na árvore especifica o valor booleano, verdadeiro ou falso, a ser retornado se essa folha for atingida (RUSSEL, 1995).

2.6 – CLASSIFICAÇÃO EM ÁRVORE DE DECISÃO

Uma árvore de decisão utiliza uma estratégia de dividir para conquistar, onde um problema complexo é decomposto em subproblemas mais simples e recursivamente a mesma estratégia é aplicada a cada subproblema (QUINLAN, 1986).

Segundo exemplo apresentado por BASGALUPP (2010), a fim de se ilustrar o funcionamento básico de uma árvore de decisão, pode-se considerar o problema para diagnóstico de pacientes de acordo com a Figura 2.11.

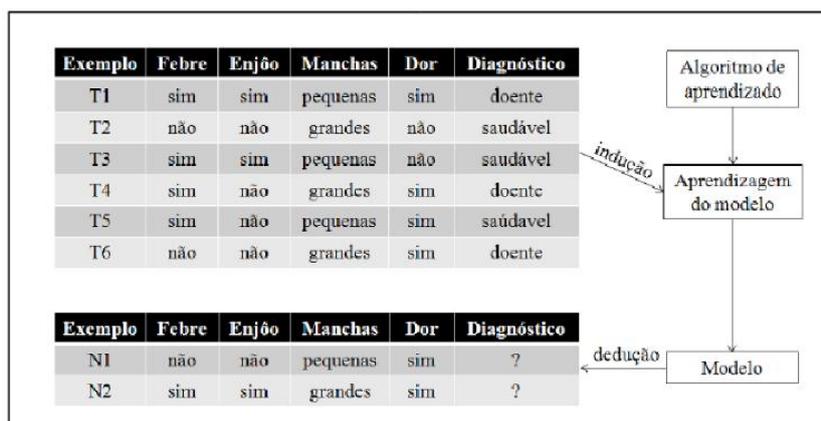


Figura 2.11 - Exemplo de dados de treinamento (*Training Data*).
 Fonte: BASGALUPP (2010).

A Figura 2.11 apresenta uma tabela com os diagnósticos, variável de saída e os sintomas, variáveis de entrada, que levaram àquele diagnóstico. Essa tabela é utilizada para que a árvore “aprenda” quais as variáveis de entrada influenciam no diagnóstico. Depois que a árvore “aprendeu” ela é utilizada para inferir o diagnóstico de saída para novos casos, com novos sintomas.

Supondo que um novo paciente chegue ao consultório médico, com a finalidade de diagnosticá-lo a primeira pergunta que pode ser feita é se tem sentido dor. A seguir, diferentes perguntas podem ser realizadas. Como por exemplo, (i) se apresenta febre ou enjoos ou (ii) ainda se tem notado alguma mancha no corpo. Isto caracteriza uma forma para solucionar um problema de classificação por intermédio de perguntas referentes a uma série de características, que no caso é o paciente. A cada pergunta respondida, outra pode ser realizada até que se chegue a uma conclusão sobre a classe a que pertence o exemplo. Tal série de perguntas e suas possíveis respostas podem ser organizadas na representação de uma árvore de decisão composta de forma hierárquica por nodos, ou nós, e arestas. Apesar dos exemplos utilizados neste texto serem de árvores binárias onde cada nó intermediário divide-se exatamente em dois nós descendentes, as árvores de decisão podem se dividir em vários nós descendentes (BASGALUPP, 2010).

Na Figura 2.12 é ilustrada uma possível árvore de decisão para o problema de diagnosticar pacientes, que podem ser classificados como “saudável” ou “doente”.

Dessa forma, torna-se possível a utilização de árvore de decisão para classificação de um paciente como saudável ou doente. Basta partir do nodo raiz da árvore e percorrê-la por meio das respostas aos testes dos nodos internos até que se chegue em um nodo folha, o qual indica a classe correspondente ao paciente. Além da obtenção da classe, a

grande vantagem que a trajetória percorrida até o nodo folha representa uma regra, facilitando a interpretação do modelo pelo usuário, no caso um médico (BASGALUPP, 2010).

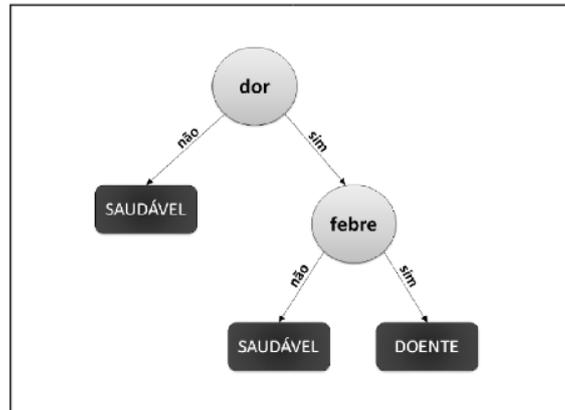


Figura 2.12 - Exemplo de árvore de decisão.
Fonte: BASGALUPP (2010).

2.7 – TÉCNICAS PARA A SELEÇÃO DE ATRIBUTOS

Uma árvore pode ser induzida pela sucessiva seleção e subdivisão de atributos. Estes atributos podem ser escolhidos de forma randômica, e eventualmente uma árvore pode ser formada de nodos terminais a partir do momento em que cada nodo possui membros de apenas uma única classe. Assim, a taxa de erro aparente é minimizada e normalmente é zero.

2.7.1 - Exemplo de indução de árvores de decisão

Segundo DURKIN (1994), indução é o processo de raciocínio sobre um dado conjunto de fatos para princípios gerais ou regras. Toma-se como exemplo a seguinte linha de pensamento: se eu dissesse a alguém que eu gosto de futebol, vôlei e natação, esta pessoa provavelmente concluiria por indução que eu gosto de esportes. A indução busca padrões em informações disponíveis com o propósito de inferir conclusões racionais.

O aprendizado indutivo tem sido uma importante área de pesquisa em Inteligência Artificial e pode ser visto como uma busca de soluções no espaço de estados do problema.

Para ilustrar o aprendizado indutivo, será apresentado um problema que procura determinar qual presente comprar para uma determinada pessoa perante algumas características (DURKIN, 1994).

O problema em questão aborda alguns pontos como: dinheiro, a idade da pessoa que receberá o presente e os tipos de presente. O problema procura achar um presente apropriado com base no dinheiro disponível e a idade da pessoa. Num primeiro momento, para facilitar a elaboração e resolução do problema, assume-se que os atributos dinheiro e idade podem assumir valores tais como: dinheiro {muito, pouco}, e idade {criança, adulto}. O conhecimento a respeito deste problema é dado por um conjunto de exemplos obtidos por um especialista no assunto, e é representado na Tabela 2.1. Os conceitos de dinheiro e idade representam os fatores de decisão ou os atributos do problema. Dos exemplos mostrados na Tabela 2.1, pode-se induzir uma busca no espaço do problema conforme a Figura 2.13.

Tabela 2.1 - Tabela de decisão para o problema do presente.

Fatores de decisão		Resultado
Dinheiro	Idade	Presente
Muito	Adulto	Carro
Muito	Criança	Computador
Pouco	Adulto	Torradeira
Pouco	Criança	Calculadora

Fonte: DURKIN (1994).

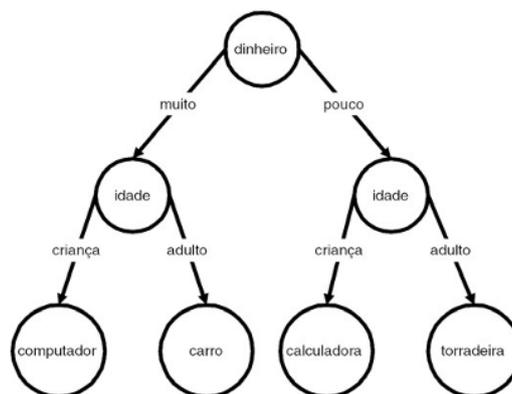


Figura 2.13 – Busca de estados do problema do presente.

Fonte: DURKIN (1994).

Por indução, pode-se criar uma árvore de decisão a partir do conjunto de exemplares para que se possa realizar o processo de busca, ou pode-se ainda criar um

conjunto de regras para se utilizar em um sistema especialista. Para o exemplo apresentado na Tabela 2.1, as regras são:

- **SE** á muito dinheiro, **E** se estiver comprando um presente para um adulto, **ENTÃO** comprar um carro;
- **SE** há muito dinheiro, **E** se estiver comprando um presente para uma criança, **ENTÃO** comprar um computador;
- **SE** há pouco dinheiro, **E** se estiver comprando um presente para um adulto, **ENTÃO** comprar uma torradeira;
- **SE** há pouco dinheiro, **E** se estiver comprando um presente para uma criança, **ENTÃO** comprar uma calculadora.

A tarefa da indução é desenvolver regras de classificação que podem determinar a classe de um objeto através dos valores de seus atributos. Os objetos são descritos em termos de uma coleção de atributos (QUINLAN, 1986). Cada atributo mede alguma característica importante do objeto; e cada objeto, no domínio da aplicação, pertence a um conjunto de classes mutuamente exclusivas onde a classe deste objeto é conhecida.

Segundo OLIVEIRA (2001), se o conjunto de treinamento contém dois objetos que têm valores idênticos para cada atributo e pertencerem a classes diferentes, é claramente impossível à diferenciação entre estes objetos com referência somente em seus dados atributos. E neste caso, os atributos serão considerados inadequados para o conjunto de treinamento e conseqüentemente para a tarefa de indução.

Um conflito ocorre quando dois exemplos contêm valores idênticos para todos os seus atributos, mas valores de classes diferentes. Um conflito normalmente significa que os atributos escolhidos são inadequados para a tarefa de classificação. Pode-se remover este problema introduzindo atributos adicionais, o que é uma tarefa para o especialista do domínio (OLIVEIRA, 2001).

As folhas da árvore de decisão são os nomes da classe, os nós representam testes baseados nos atributos com ramos rotulados, com os possíveis valores do atributo, para um resultado de classificação. Para classificar um objeto, começa-se da raiz da árvore, avalia-se o teste, ou seja, o nó da árvore é comparado com o respectivo atributo do objeto em questão, partindo pelo ramo determinado pelo valor do atributo do objeto que se pretende classificar, e o processo continua até que uma folha seja encontrada, na qual o objeto é afirmado a pertencer à classe nomeada pela folha.

O principal propósito da indução é construir árvores de decisão que possam classificar de forma correta os exemplares do conjunto de treinamento, mas também o restante dos exemplares que compõe o conjunto de dados.

2.8 – ALGORITMO ID3

Em meados da década de 70, um pesquisador em inteligência artificial, chamado J. Ross Quinlan, da Universidade de Sydney, Austrália, utilizou o modelo de formação de conceitos para desenvolver um programa chamado ID3 (*Itemized Dichotomizer 3*). O algoritmo desenvolvido por Quinlan “aprendeu”, através de um pequeno conjunto de treinamento, como organizar e processar um amplo conjunto de dados. Ele utilizou a estratégia de dividir para conquistar, combinado com a lógica dicotômica para produzir bons resultados em relação ao baixo tempo de processamento (GESTWICKI, 1997).

O ID3 é um algoritmo que procura usar a lógica e a matemática para processar, organizar e simplificar grandes conjuntos de dados. O algoritmo ID3 é descendente do *Concept Learning System* (CLS) e constrói árvores de decisão de maneira *top down*. Sua metodologia de geração de árvore consiste em começar pelo atributo mais relevante e continuar pelos outros atributos segundo a avaliação da entropia de cada um. O algoritmo visa resolver problemas que tenham atributos categóricos (não contínuos) e sem ruídos. Sendo assim, atributos ruidosos devem ser tratados previamente (HOLSHEIMER e SIEBES, 1991).

De acordo com OLIVEIRA (2001), para a formação da árvore de decisão o conjunto de treinamento T é dividido em subconjuntos T_i onde esta divisão é feita de acordo com o atributo A , escolhido para raiz da árvore, e seus possíveis valores distintos do atributo; cada T_i conterá apenas os objetos de T com valores v_i de A e consequentemente cada T_i será menor que T . Sendo que, a partir do atributo raiz A partem os ramos rotulados com os possíveis valores distintos de A para cada subconjunto T_i . A ideia básica é “dividir e conquistar”, isto é, proceder com esta escolha da raiz e consequentemente com a divisão do conjunto em subconjuntos para cada T_i , um por vez (estratégia de busca *top down* e *hill climber*), até encontrar uma folha para este. O resultado final será uma árvore para T , pois esta estratégia renderá subconjuntos até que satisfaçam a exigência de uma classe para uma folha, ou seja, até que nestes subconjuntos tenham objetos pertencentes apenas a uma única classe, logo uma folha é encontrada.

Dado um conjunto de exemplares, cada exemplar possui a mesma estrutura, consistindo em um número pré-determinado de atributos. Um destes atributos representa a categoria do registro/exemplar. O problema é determinar a árvore de decisão que, com base nas respostas e perguntas com relação aos atributos não categóricos, prediz corretamente o valor do atributo categórico. Normalmente o atributo categórico assume apenas valores {verdadeiro, falso}, ou {sucesso, falha}, ou algo equivalente.

Na Tabela 2.2, há registros que informam as condições climáticas para a realização de um jogo de golfe. O atributo categórico Classe especifica se deverá haver ou não o jogo, baseado nos atributos não categóricos Céu, Temperatura, Umidade e Vento.

Tabela 2.2 – Condições para a realização do jogo.

Nº exemplar	Céu	Temperatura	Umidade	Vento	Classe
1	Sol	Alta	Alta	Não	Não joga
2	Sol	Alta	Alta	Sim	Não joga
3	Nublado	Alta	Alta	Não	Joga
4	Chuva	Alta	Alta	Não	Joga
5	Chuva	Baixa	Normal	Não	Joga
6	Chuva	Baixa	Normal	Sim	Não joga
7	Nublado	Baixa	Normal	Sim	Joga
8	Sol	Suave	Alta	Não	Não joga
9	Sol	Baixa	Normal	Não	Joga
10	Chuva	Suave	Normal	Não	Joga
11	Sol	Suave	Normal	Sim	Joga
12	Nublado	Suave	Alta	Sim	Joga
13	Nublado	Alta	Normal	Não	Joga
14	Chuva	Suave	Alta	Sim	Não joga

Fonte: QUINLAN (1986).

Para a criação da árvore o algoritmo ID3 calcula a Entropia do conjunto de atributos. A entropia do sistema representa o quanto de informação um conjunto de dados carrega, ou seja, a Entropia mede o grau de “impureza” (desordem) dos dados.

Se a entropia = 0, o conjunto de dados é “puro”, se entropia = 1, o conjunto é totalmente “impuro”.

Por exemplo, para problemas com duas classes, tipo: Sim e Não, se um conjunto tiver apenas exemplos da classe **Sim** ou da classe **Não**, a sua entropia será **0**, se um conjunto contiver o mesmo número de exemplos da classe **Sim** e da classe **Não**, a sua entropia será **1**.

A entropia é calculada para cada atributo (variável de entrada) do conjunto de dados e leva em conta o número de classes que cada atributo possui, então o atributo com a menor entropia é escolhida para ser o nó raiz, o atributo com a segunda menor entropia será o segundo nó e assim por diante.

Para um atributo com n classes, a entropia é definida pela Eq. (2.4):

$$Entropia = \sum_{i=1}^n - p_i \log_2 p_i \quad (2.4)$$

Sendo:

- p_i = a probabilidade da classe ocorrer dentro daquele atributo;
- \log_2 = o logaritmo na base 2 da probabilidade p_i .

É feito o somatório negativo do produto de p_i e o logaritmo. Para ilustrar melhor observa-se o seguinte exemplo:

M representa um atributo de um sistema em estudo, que possui apenas dois tipos de ocorrência (classes), o conjunto $\{1, 2\}$. O conjunto de dados de M possui dois valores distintos, onde a probabilidade de se encontrar cada um dos valores é igual a meio ($1/2$), a entropia é calculada aplicando-se a Eq. (2.4):

$$Entropia (M) = - 0,5 * \log_2(0,5) - 0,5 * \log_2(0,5)$$

Desta maneira, o ID3 na concepção da árvore de decisão escolhe o atributo (céu) de maior ganho de informação para ser o nodo raiz da árvore, mostrado na Figura 2.14.

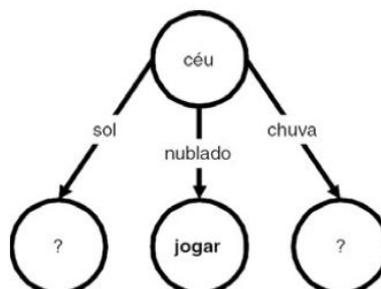


Figura 2.14 – Construção da árvore em nível parcial.
Fonte: QUINLAN (1986).

Analisando-se a Figura 2.14, verifica-se que os ramos sol e chuva ainda estão indefinidos, e o processo deve continuar no próximo nível da árvore.

A Figura 2.15 apresenta a árvore de decisão final gerada pelo ID3 para o conjunto de treinamento da Tabela 2.2, os detalhes de cálculo foram suprimidos.

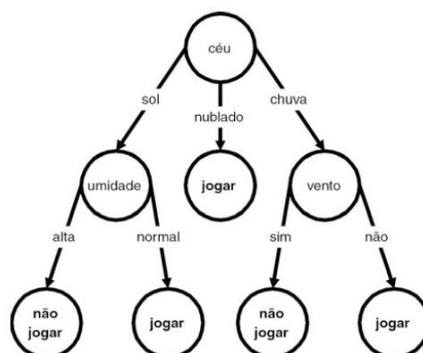


Figura 2.15 – Árvore de decisão final.
Fonte: QUINLAN (1986).

Observa-se que o atributo temperatura não foi selecionado para fazer parte da árvore, devido ao fato do ID3 o ter considerado irrelevante para tarefa de classificação.

2.9 – ALGORITMO C4.5

A necessidade de se trabalhar com atributos do tipo contínuo fez com que sucessivas melhorias fossem desenvolvidas no algoritmo ID3. Em 1993, J. Ross Quinlan apresenta seu trabalho intitulado "C4.5: Programs for machine learning " (QUINLAN, 1993), que foi um aprimoramento do algoritmo ID3, tornando possível trabalhar com atributos contínuos.

Dentre os vários algoritmos de indução de árvores de decisão, o ID3 e seu sucessor C4.5 são os mais populares na comunidade científica. Estes algoritmos e suas variações são tema de inúmeros artigos de pesquisa desde que Quinlan apresentou o ID3.

O C4.5 gera um classificador que é capaz de agir como um especialista, classificando inclusive casos desconhecidos. Este foi desenvolvido com o intuito de tornar o modelo de classificação mais inteligível (POZO, 2002).

Assim como o ID3, o C4.5 constrói a árvore de decisão através do conjunto de treinamento, ou de exemplos, combinando uma estrutura de dados em árvore que pode ser usada para classificar novos exemplos. O C4.5 também emprega os uso da Teoria da Informação para avaliar a qualidade de um nodo de teste (SEGRE, 1993). O algoritmo de uma forma peculiar extrai a máxima quantidade de informação de um conjunto de

exemplares dada a condição de que somente um atributo será utilizado para realização do teste.

Este novo procedimento de indução de árvores de decisão apresenta uma série de extensões com relação ao algoritmo original ID3.

O C4.5 representa o resultado de vários anos de investigação na aprendizagem automática e na extração do conhecimento, sendo tomado como ponto de referência para o desenvolvimento de novos algoritmos; pois apresenta resultados que demonstram que este procedimento de indução de árvores de decisão oferece uma boa precisão na classificação e é considerado um dos mais rápidos (MICHIE *et al.*, 1994).

Na construção da árvore de decisão o C4.5 pode lidar com conjuntos de treinamentos que possuem exemplares com valores de atributos desconhecidos, avaliando o ganho de um atributo considerando apenas os registros que possuem atributos definidos. Na construção da árvore de decisão, podem ser classificados registros que possuem valores de atributos desconhecidos através da estimativa de vários resultados. O C4.5 também trata o caso de atributos contínuos, cujo domínio pertence ao conjunto dos números reais.

2.10 – ALGORITMO J48

O algoritmo J48 surgiu da necessidade de recodificar o algoritmo C4.5, que, originalmente, é escrito na linguagem C, para a linguagem Java. Ele tem a finalidade de gerar uma árvore de decisão baseada em um conjunto de dados de treinamento, sendo este modelo usado para classificar as instâncias no conjunto de teste. Um dos aspectos para a grande utilização do algoritmo J48 pelos especialistas em Data Mining é que o mesmo se mostra adequado para os procedimentos, envolvendo as variáveis (dados) qualitativas contínuas e discretas presentes nas bases de dados.

O algoritmo J48, proposto por QUINLAN (1993), é considerado o que apresenta o melhor resultado na montagem de árvores de decisão, a partir de um conjunto de dados de treinamento. Para a montagem da árvore, o algoritmo J48 utiliza a abordagem de dividir-para-conquistar, onde um problema complexo é decomposto em subproblemas mais simples, aplicando recursivamente a mesma estratégia a cada subproblema, dividindo o espaço definido pelos atributos em subespaços, associando-se a eles uma classe (QUINLAN, 1993).

2.11 – SOFTWARE WEKA

O software utilizado neste trabalho para criar uma árvore de decisão, para a simplificação das regras Fuzzy chama-se WEKA (Waikato Environment for Knowledge Analysis – Ambiente Waikato para Análise do Conhecimento), desenvolvido pela Universidade de Waikato, Nova Zelândia.

O software WEKA é um software livre (de código aberto) para mineração de dados e geração de Árvores de Decisão, foi desenvolvido em Java, dentro das especificações da GNU (General Public License).

O WEKA possui dezenas de algoritmos de classificação (mais de 50), incluindo o ID3, C4.5 e J48, além de vários para agrupamentos e alguns de regras de associação, além de seleção de atributos. Ainda, existem dezenas de funções dos chamados filtros, que nada mais são do que algoritmos de tratamento e transformação de dados.

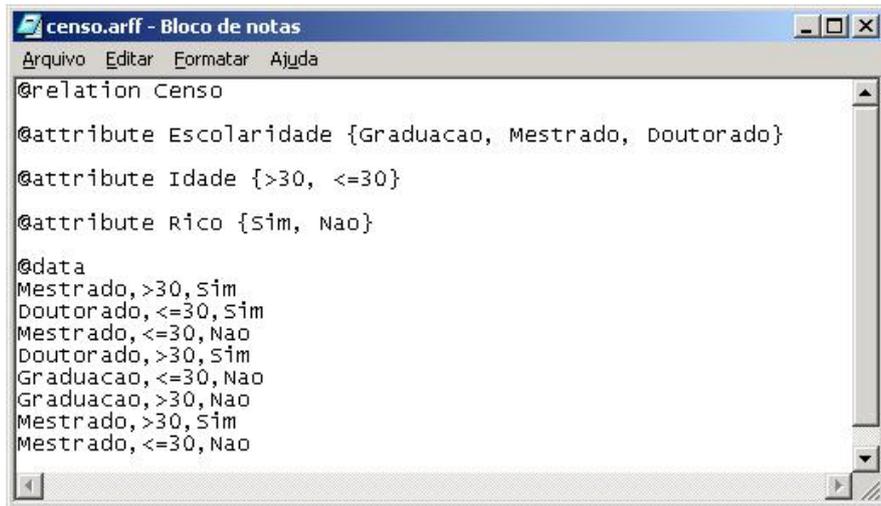
2.11.1 – Árvore de decisão

A ferramenta WEKA possui como ponto forte a extração de classificadores em bases de dados. Um classificador (ou modelo de classificação) é utilizado para identificar a classe à qual pertence uma determinada observação de uma base de dados, a partir de suas características (seus atributos).

A mineração de modelos de classificação em bases de dados é um processo composto por duas fases: aprendizado e teste. Na fase de aprendizado, um algoritmo classificador é aplicado sobre um conjunto de dados de treinamento. Como resultado, obtêm-se a construção do classificador propriamente dito. Tipicamente, o conjunto de treinamento corresponde a um subconjunto de observações selecionadas de maneira aleatória a partir da base de dados que se deseja analisar. Cada observação do conjunto de treinamento é caracterizada por dois tipos de atributo: o atributo classe, que indica a classe à qual a observação pertence; e os atributos preditivos, cujos valores serão analisados para que seja descoberto o modo como eles se relacionam com o atributo classe.

A ferramenta WEKA trabalha com arquivos de entrada no formato ARFF (Attribute-Relation File Format - Formato de Arquivo de Relação de Atributos), que corresponde a um arquivo texto contendo um conjunto de observações, precedido por um pequeno cabeçalho. O cabeçalho é utilizado para fornecer informações a respeito dos

campos que compõem o conjunto de observações. Dessa forma, antes da mineração de dados, a ferramenta pode verificar alguma inconsistência na base de dados e sinalizá-la. A Figura 2.16 ilustra um exemplo de um arquivo ARFF, contendo um cabeçalho e um conjunto de 8 registros que representam a base de dados apresentada na Tabela 2.3.



```
@relation Censo
@attribute Escolaridade {Graduacao, Mestrado, Doutorado}
@attribute Idade {>30, <=30}
@attribute Rico {sim, Nao}
@data
Mestrado,>30,sim
Doutorado,<=30,sim
Mestrado,<=30,Nao
Doutorado,>30,sim
Graduacao,<=30,Nao
Graduacao,>30,Nao
Mestrado,>30,sim
Mestrado,<=30,Nao
```

Figura 2.16 - Arquivo ARFF.

Observe na Figura 2.16 que o cabeçalho contém a declaração da relação que o arquivo representa (comando `@relation`), uma lista de atributos (comando `@attribute`) e a relação de valores que os mesmos podem assumir. O conjunto de observações é precedido por um comando `@data`. Cada observação é representada por uma linha. Os valores dos campos dentro de uma observação devem ser separados utilizando a vírgula.

A representação da idade mostra os valores que este atributo pode ter (>30 e <=30).

Tabela 2.3 - Base de Dados.

NOME	ESCOLARIDADE	IDADE	RICO <i>(atributo classe)</i>
Alva	Mestrado	>30	Sim
Amanda	Doutorado	<=30	Sim
Ana	Mestrado	<=30	Não
Eduardo	Doutorado	>30	Sim
Inês	Graduação	<=30	Não
Joaquim	Graduação	>30	Não
Maria	Mestrado	>30	Sim
Raphael	Mestrado	<=30	Não

O instalador da ferramenta WEKA pode ser obtido de maneira gratuita (juntamente com seu código fonte) no site <http://www.cs.waikato.ac.nz/~ml/weka>. Uma vez instalado, o sistema WEKA pode ser utilizado para minerar árvores de decisão através da execução dos seguintes passos:

PASSO 1: Executar o programa. A partir do menu Iniciar / Programas, selecione WEKA e clique em WEKA. O menu principal WEKA GUI Chooser será exibido na tela. Clique no botão “Explorer”, mostrado na Figura 2.17.



Figura 2.17 - WEKA GUI Chooser.

PASSO 2: Importar o arquivo ARFF. Após iniciar o WEKA Explorer, a opção “Open File” deve ser utilizada para abrir o arquivo ARFF desejado.

PASSO 3: Selecionar os Atributos. Em seguida, o WEKA abrirá uma tela que permite com que o usuário possa definir qual o atributo da base que será utilizado como classe e quais os atributos que serão utilizados como preditivos, mostrado na Figura 2.18.

No momento da importação, por *default*, o WEKA irá considerar o último atributo especificado no cabeçalho do arquivo ARFF, como o atributo classe, enquanto os demais atributos serão tratados como atributos preditivos. Observe que na tela da Figura 2.18, também é possível consultar gráficos de barra que indicam os cruzamentos de frequência envolvendo todos os atributos preditivos e o atributo classe.

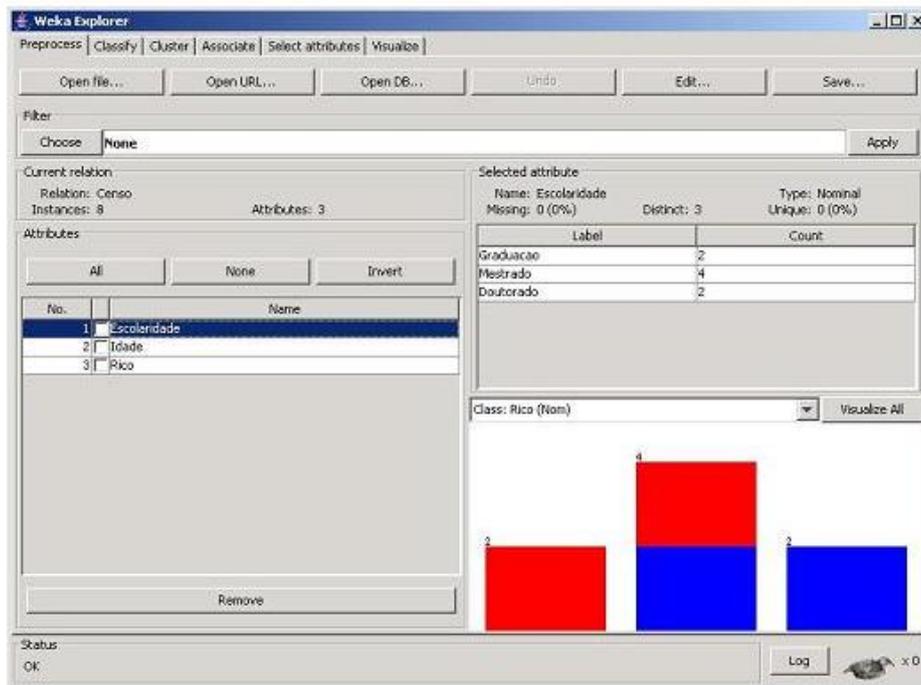


Figura 2.18 - Seleção da classe e dos atributos preditivos.

PASSO 4: Selecionar o Algoritmo de Mineração. Clique na aba “*Classify*”. A partir desta tela é possível escolher e executar um algoritmo de classificação sobre a base de dados importada. Os resultados da mineração também poderão ser consultados neste mesmo local. Clique no botão “*Choose*”. Será aberta uma janela que permitirá a escolha do algoritmo de mineração de dados. Como vão ser utilizados atributos categóricos (classe), clique na pasta “*trees*” (algoritmos de árvore de decisão) e selecione a opção “*Id3*”, mostrado na Figura 2.19.

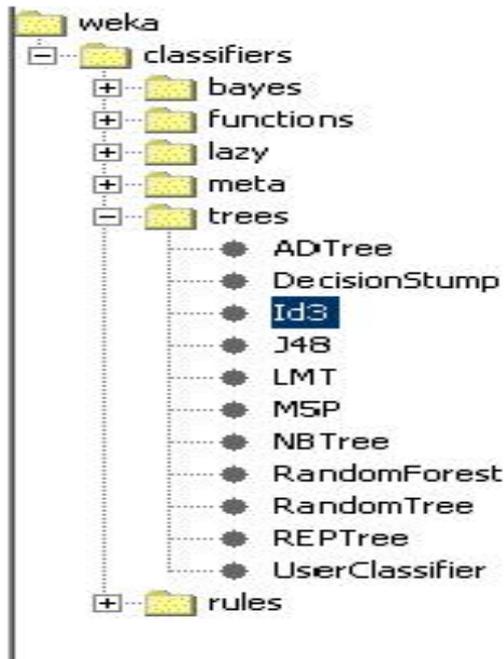


Figura 2.19 - Seleção do algoritmo de mineração de dados.

PASSO 5: Executar o Algoritmo de Mineração. No painel “Test options” selecione a opção “Use training set”. Esta seleção indica ao WEKA que toda a base de dados será utilizada como base de treinamento durante o processo de mineração. A seguir clique no botão "Start". Os dados da árvore de decisão gerada pelo algoritmo ID3, são apresentados no canto direito da tela do WEKA, conforme ilustra a área destacada no círculo vermelho da Figura 2.20. Na mesma tela são apresentadas algumas medidas de interesse que indicam a qualidade da árvore minerada. Essas informações serão melhor detalhadas no capítulo 3.

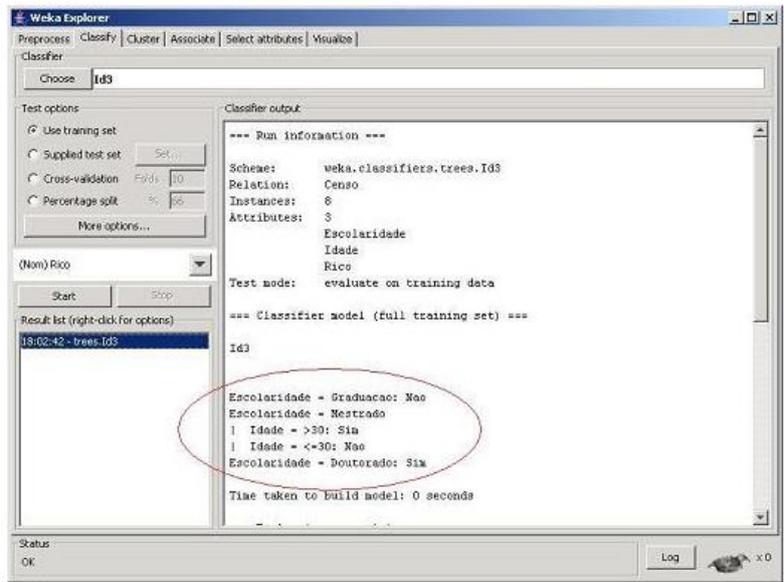


Figura 2.20 - Árvore de decisão minerada pelo WEKA.

CAPÍTULO 3

MATERIAIS E MÉTODOS

3.1 – METODOLOGIA EXPERIMENTAL

Belém é o município mais populoso do Pará e o segundo da região Norte, com uma população de 1.446.042 habitantes, segundo estimativa do IBGE, em 2016, e o 12º município mais populoso do Brasil. Ocupa a 22ª posição no ranking de IDH por capital (0,746, alto) e a sexta posição na lista de maiores IDH da região Norte – 3º maior IDH por capital por região.

Ananindeua é um município da Região Metropolitana de Belém. É o segundo município mais populoso do estado e o quarto da Região Norte do Brasil. Sua população é estimada de 510.830 habitantes, segundo dados do IBGE 2016, superada por Manaus, Belém e Porto Velho (G1 PARÁ, 2016). A localização dos municípios é mostrada na Figura 3.1.



Figura 3.1 – Localização dos municípios de Belém e Ananindeua.

Estes dois municípios do Pará vêm sofrendo uma grande verticalização nos últimos 30 anos, produzindo uma demanda imobiliária crescente a cada ano.

3.2 - CRIAÇÃO DO MODELO FUZZY

Para a elaboração do modelo Fuzzy, foi utilizado um levantamento de campo realizado pela Faculdade de Engenharia Civil/ITEC/UFPA, para uma atividade específica, totalizando 99 apartamentos, nos municípios de Ananindeua e Belém, cujos dados foram cedidos para a realização deste trabalho e uma parte deles está mostrada na Figura 3.2. A listagem completa do arquivo encontra-se nos Anexos A e B.

Item	FOTO/AMOSTRA	ENDEREÇO	CIDADE	VAGAS DE GARAGEM	ÁREA PRIVATIVA	Padrão	IDADE APARENTE	LOCAL	VALOR m ²
1		Rua Boaventura da Silva, 1227, Unid 504	1	2	140,45	5	5	4	5.339,98
2		Rua dos Mundurucus, nº 1581, Apto 1801	1	2	139,48	5	1	4	6.094,06
3		Av. Av. Gentil Bittencourt, nº 1014, apto 2301	1	2	354,40	4	1	5	5.135,44
4		Av. Vinte e Cinco de Setembro, nº 1687-1695, Apto 1604	1	2	92,51	3	1	5	4.864,34

Figura 3.2 – Amostra dos dados levantados pela FEC/ITEC/UFPA.
Fonte: FEC/ITEC/UFPA.

Os atributos utilizados na elaboração do modelo, a partir do levantamento da Figura 3.2, foram os seguintes:

- **Área Privativa** – variável quantitativa contínua, definindo a área útil do imóvel em m², considerando o intervalo de 0 m² a 500 m²;
- **Padrão** – variável qualitativa, atribuindo valores discretos de 0 a 5, para o acabamento do imóvel;
- **Idade** – variável qualitativa, atribuindo valores discretos de 0 a 40 para o tempo de construção do imóvel;
- **Local** – variável qualitativa, atribuindo valores discretos de 0 a 5 para o padrão de localização do imóvel na cidade;
- **Cidade** – variável qualitativa dicotômica, atribuindo o valor 0 para imóvel situado no município de Ananindeua e 1 para imóvel situado no município de Belém, que foram os dois municípios onde foi realizado o estudo.
- **Vagas** – variável quantitativa, atribuindo valores discretos de 0 a 4 para o número de vagas de garagem oferecidas pelo imóvel.

A variável de saída do modelo, ou a variável dependente foi definida como sendo o **Valor do m²** (em R\$/m²) do imóvel, com intervalo contínuo entre R\$ 0,00/m² a R\$ 9.000,00/m².

Assim, a relação final das variáveis linguísticas de entrada e saída, adotadas, está relacionada na Tabela 3.1 e suas funções de pertinência foram definidas baseadas em DUARTE (2006) e em pesquisa com especialistas na área.

Tabela 3.1 – Variáveis linguísticas e suas funções de pertinência.

Variáveis Linguísticas	Funções de Pertinência
Área	pequena
	média
	grande
Padrão	inferior
	médio
	superior
Idade	novo
	médio
	antigo
Local	inferior
	médio/inferior
	médio
	médio/superior
	superior
Cidade	Ananindeua
	Belém
Vagas	inferior
	médio
	superior
Valor (R\$/m ²)	baixo
	médio
	alto
	muito_alto
	muitíssimo_alto

A Tabela 3.1 mostra as Variáveis Linguísticas de entrada (Área, Padrão, Idade, Local, Cidade e Vagas) e de saída [Valor (R\$/m²)], com suas respectivas Funções de Transferência. A definição das funções de transferência foi feita tomando como base DUARTE (2006) e consultas com especialistas na área.

Para a criação do sistema Fuzzy, foi utilizado o software MatLab, no qual as variáveis linguísticas e suas funções de pertinência tiveram as definições mostradas nas Figuras 3.3 a 3.9:

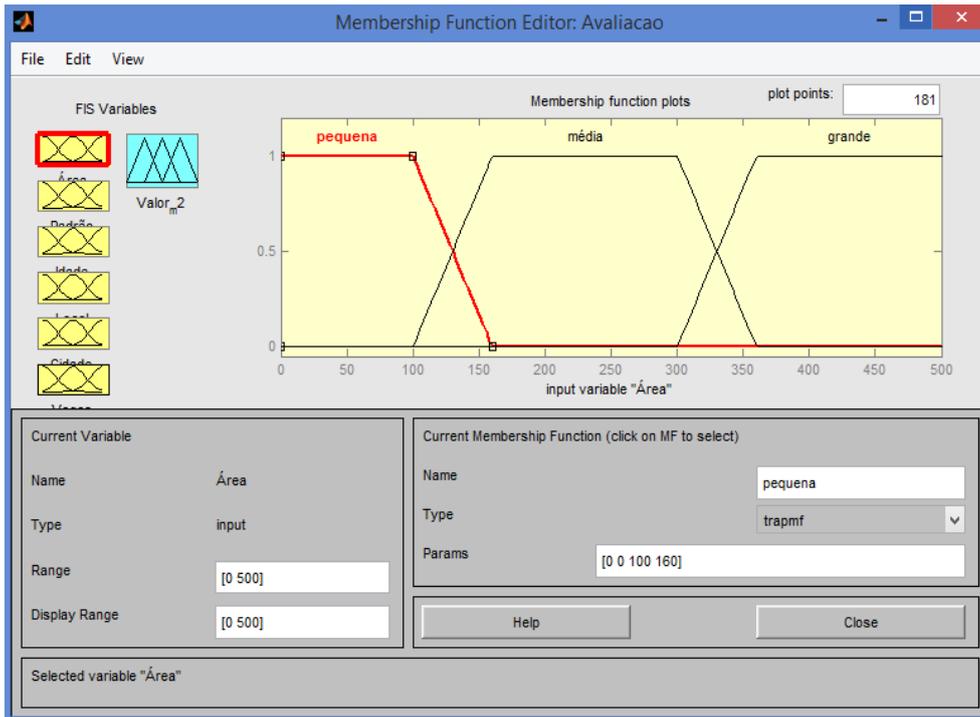


Figura 3.3 – Variável linguística de entrada - Área.

Na Figura 3.3 foi feita a definição da variável linguística Área e suas funções de pertinência, de acordo com a definição da Tabela 3.1. Observa-se que a distribuição das funções de pertinência Pequena, Média e Grande, foi feita no intervalo de 0 a 500, tendo como base DUARTE (2006) e consultas com especialistas na área.

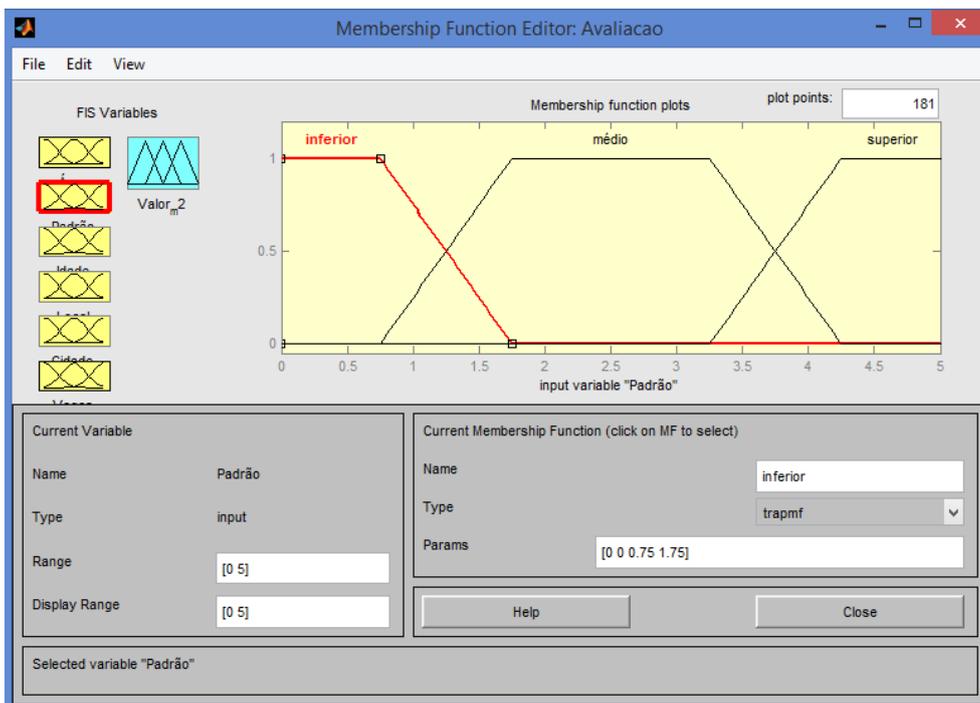


Figura 3.4 – Variável linguística de entrada – Padrão.

Na Figura 3.4 foi feita a definição da variável linguística Padrão e suas funções de pertinência, de acordo com a definição da Tabela 3.1. Observa-se que a distribuição das funções de pertinência Inferior, Médio e Superior, foi feita no intervalo de 0 a 5, tendo como base DUARTE (2006) e consultas com especialistas na área.

No modelo criado esta variável linguística pode assumir valores contínuos, mas no levantamento mostrado no Anexo A foram registrados somente valores discretos.

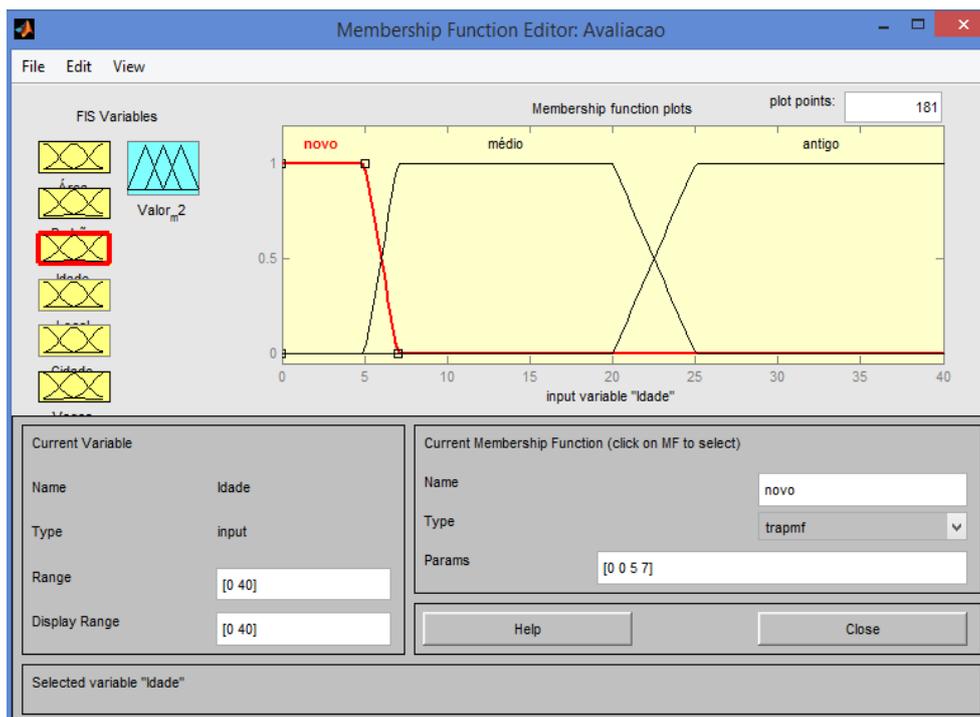


Figura 3.5 – Variável linguística de entrada - Idade.

Na Figura 3.5 foi feita a definição da variável linguística Idade e suas funções de pertinência, de acordo com a definição da Tabela 3.1. Observa-se que a distribuição das funções de pertinência Novo, Médio e Antigo, foi feita no intervalo de 0 a 40, tendo como base DUARTE (2006) e consultas com especialistas na área.

No modelo criado esta variável linguística pode assumir valores contínuos, mas no levantamento mostrado no Anexo A foram registrados somente valores discretos.

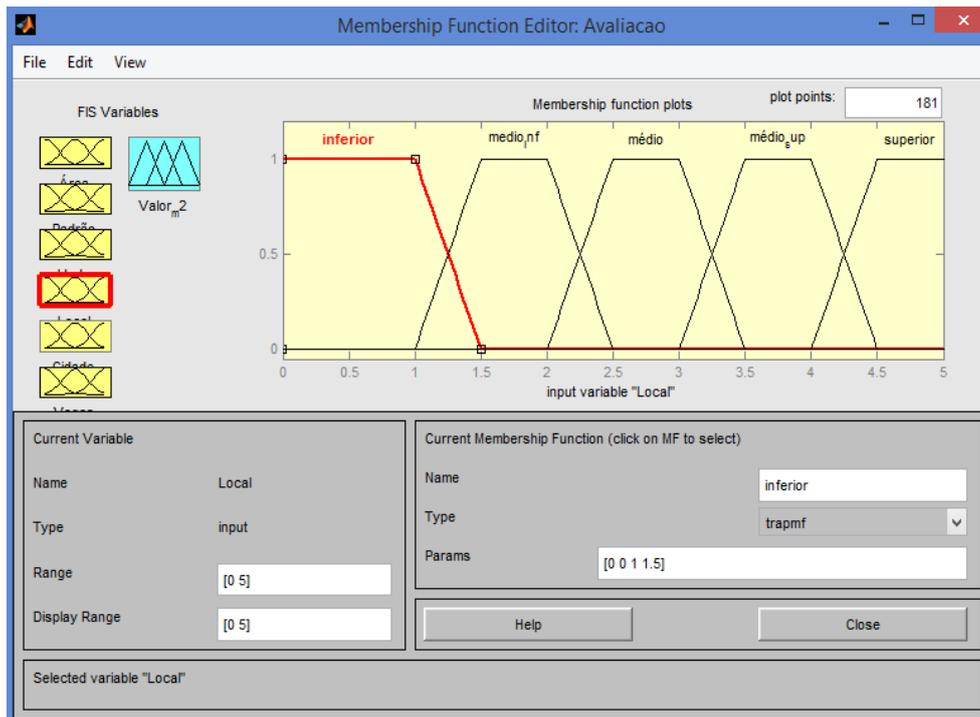


Figura 3.6 – Variável linguística de entrada - Local.

Na Figura 3.6 foi feita a definição da variável linguística Local e suas funções de pertinência, de acordo com a definição da Tabela 3.1. Observa-se que a distribuição das funções de pertinência Inferior, Médio/Inferior, Médio, Médio/Superior e Superior, foi feita no intervalo de 0 a 5, tendo como base DUARTE (2006) e consultas com especialistas na área.

No modelo criado esta variável linguística pode assumir valores contínuos, mas no levantamento mostrado no Anexo A foram registrados somente valores discretos.

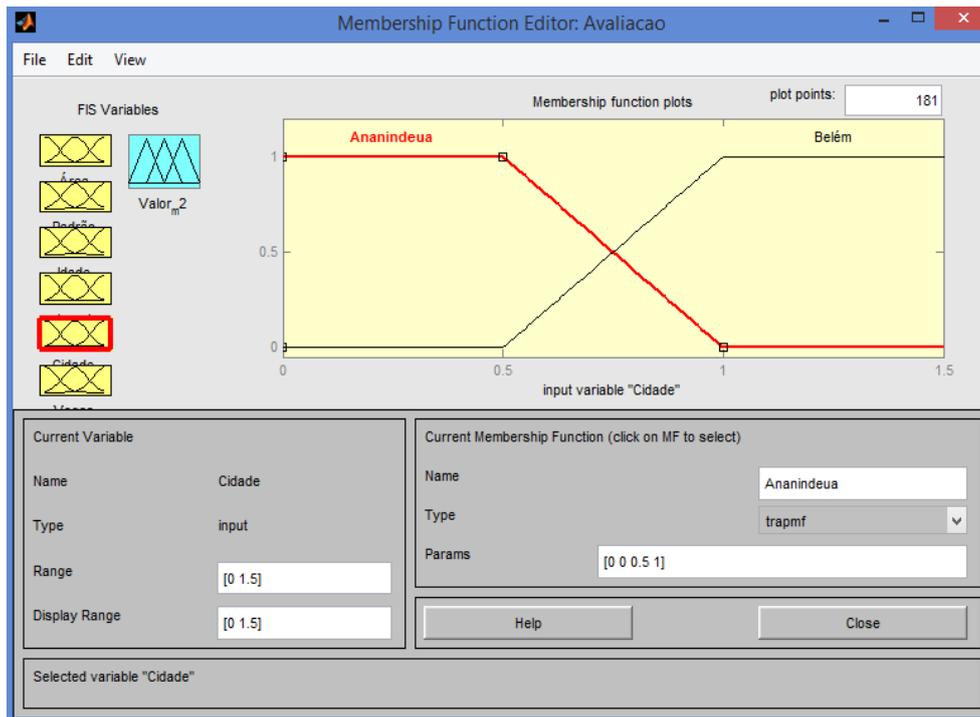


Figura 3.7 – Variável linguística de entrada - Cidade.

Na Figura 3.7 foi feita a definição da variável linguística Cidade e suas funções de pertinência, de acordo com a definição da Tabela 3.1. Observa-se que a distribuição das funções de pertinência Ananindeua e Belém, foi feita no intervalo de 0 a 1,5, isto para podermos definir a função trapezoidal, mas os valores assumidos só podem ser 0 ou 1, apesar de podermos estender seu tipo para valores contínuos e tratar os casos dos imóveis na fronteira entre Belém e Ananindeua. Estas definições tiveram como base DUARTE (2006) e consultas com especialistas na área.

No modelo criado esta variável linguística pode assumir valores contínuos, imóveis localizados nos limites dos municípios, mas no levantamento mostrado no Anexo A foram registrados somente valores discretos.

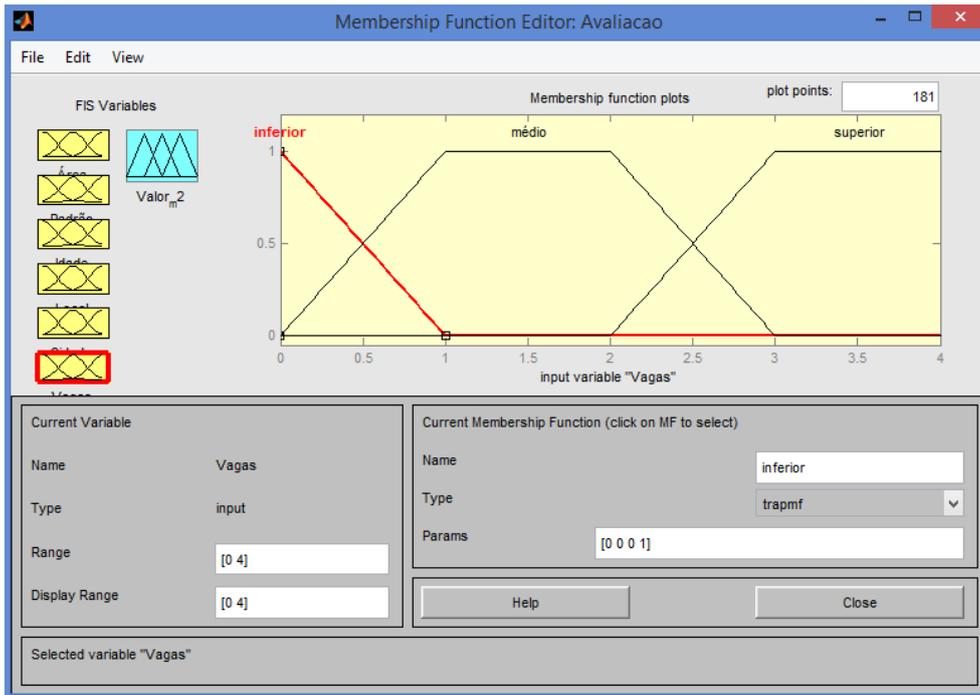


Figura 3.8 – Variável linguística de entrada - Vagas.

Na Figura 3.8 foi feita a definição da variável linguística Vagas e suas funções de pertinência, de acordo com a definição da Tabela 3.1. Observa-se que a distribuição das funções de pertinência Inferior, Médio e Superior, foi feita no intervalo de 0 a 4, tendo como base DUARTE (2006) e consultas com especialistas na área.

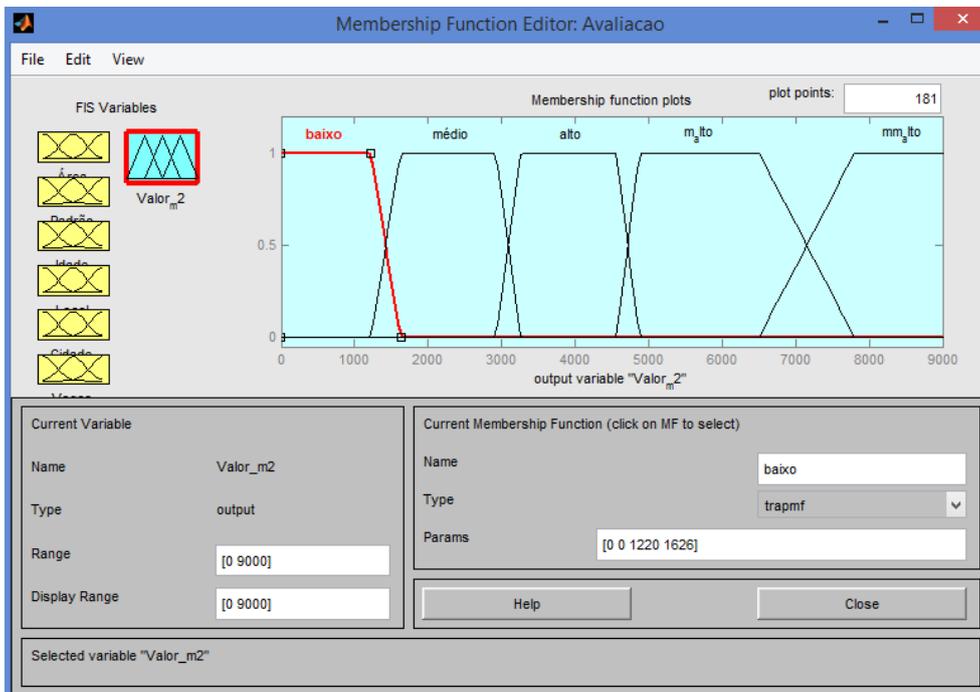


Figura 3.9 – Variável linguística de saída – Valor_m².

Na Figura 3.9 foi feita a definição da variável linguística Valor_m2 e suas funções de pertinência, de acordo com a definição da Tabela 3.1. Observa-se que a distribuição das funções de pertinência Baixo, Médio, Alto, Muito Alto e MUITÍSSIMO Alto, foi feita no intervalo de 0 a 9000, tendo como base DUARTE (2006) e consultas com especialistas na área.

O número de funções de pertinência, de cada variável linguística de entrada mostradas na Tabela 3.1, multiplicadas resultam no número de Regras Fuzzy que irá compor a Base de Regras do modelo proposto. Desta forma este cálculo resultaria em um total de **810 regras** ($3 \times 3 \times 3 \times 5 \times 2 \times 3$).

Com esse número de regras o modelo proposto torna-se, operacionalmente, difícil de implementar.

Desta forma, foi feita uma pesquisa para a proposta de uma nova forma de otimização deste procedimento.

3.3 – CRIAÇÃO DA ÁRVORE DE DECISÃO

Para a otimização do processo de criação das regras Fuzzy, foram utilizados, como referência, os trabalhos realizados em (OLARU, 2003), onde os autores fizeram a proposta de um método chamado *soft decision trees* (SDT), uma variante da aprendizagem indutiva da árvore de decisão clássica, para um sistema Fuzzy de análise de segurança de um sistema de energia elétrica, onde as entradas foram dois atributos que caracterizam o estado do sistema, denotados respectivamente por "Pu" e "Qu" (potência ativa e reativa, respectivamente) e a saída numérica que mostra o nível de segurança do sistema. As árvores predizem o grau de adesão dos dados numéricos de saída para a classe difusa, como uma forma de ratificar a defuzzyficação.

Esta foi uma aplicação muito específica, mas serviu como modelo para a utilização conjunta das duas técnicas Fuzzy e árvore de decisão, utilizadas neste trabalho.

Conforme demonstrado na seção 2.7.1, por indução, pode-se criar uma árvore de decisão a partir do conjunto de exemplares para que se possa realizar o processo de busca, ou pode-se ainda criar um conjunto de regras para se utilizar em um sistema especialista.

Neste trabalho, foi criado um conjunto de regras para ser utilizado no modelo Fuzzy proposto, em substituição ao conjunto de 810 regras originalmente calculado.

Basicamente, a ideia consiste na submissão dos dados de entrada do modelo Fuzzy, como dados de treino do software WEKA, para a criação de uma árvore de

decisão, tendo como Atributos os parâmetros do imóvel e como Classe o Valor Unitário do metro quadrado. A partir da árvore gerada, os seus nós e ramos são percorridos para se obter as regras de classificação, que serão utilizadas como regras do modelo Fuzzy proposto.

O conjunto de dados utilizados para a criação da árvore é composto de 75 registros, oriundos do levantamento feito pela FEC/ITEC/UFPA, mostrado parcialmente na Tabela 3.2. A listagem completa deste conjunto é mostrada no Apêndice A.

Os Atributos correspondem aos dados: área privativa, padrão, idade, local, cidade e vagas, que serão analisados pela Árvore. A Classe corresponde ao resultado final que se deseja obter é o Valor Unitário.

No processo de criação da árvore é feito um treinamento, com uma parte de um conjunto de dados previamente classificados e um teste, com uma outra parte do mesmo conjunto de dados, não utilizada no treino (ARTERO, 2009).

O conjunto de dados de treino utilizado correspondeu a 70% dos 75 registros do conjunto das informações mostradas na Tabela 3.2 e o de teste aos outros 30%. A seleção de quais dados serão de treino ou teste é feita de forma aleatória, pelo software gerador da árvore de decisão.

Para a criação do conjunto de treino, a Classe teve de ser adequada, o Valor Unitário é um dado contínuo, numérico, a árvore deve classificar um atributo categórico, não contínuo, discreto, que vai significar a variável linguística para este dado, que, de acordo com a Tabela 3.1, corresponde às funções de pertinência: Baixo, Médio, Alto, Muito Alto e MUITÍSSIMO Alto. O critério adotado para fazer a conversão para as funções de pertinência definidas, foi utilizar os valores numéricos e suas funções de pertinências correspondentes, mostrado na Figura 3.9.

Esta atribuição dos valores das classes da variável dependente Valor Unitário, mostrada na Tabela 3.2, foi feita com a colaboração de especialistas em avaliação de imóveis.

Parte do arquivo de treino/teste é mostrado na Tabela 3.2 e o conjunto total está no Apêndice A.

Tabela 3.2 – Conjunto de treino da árvore de decisão.

Área	Padrão	Idade	Local	Cidade	Vagas	Valor
20	3	3	3	1	0	mm_alto
30	1	35	1	1	0	alto
40	1	30	5	1	0	alto
43	1	25	4	1	0	alto
44	3	1	3	1	1	m_alto
54	2	2	3	1	1	m_alto
56	3	2	3	1	1	m_alto
59	2	1	2	0	2	m_alto
60	2	1	3	1	1	m_alto
63	1	1	1	1	1	alto
65	1	5	2	1	1	alto
67	1	1	3	1	1	alto
69	3	3	2	1	1	alto
72	3	5	5	1	1	m_alto

3.3.1 - Utilização da Ferramenta WEKA

Para a criação da Árvore de Decisão, a partir do conjunto de treino, foi utilizada a ferramenta de mineração WEKA. Após a leitura da base de dados de treino/teste, o software apresenta a tela mostrada na Figura 3.10.

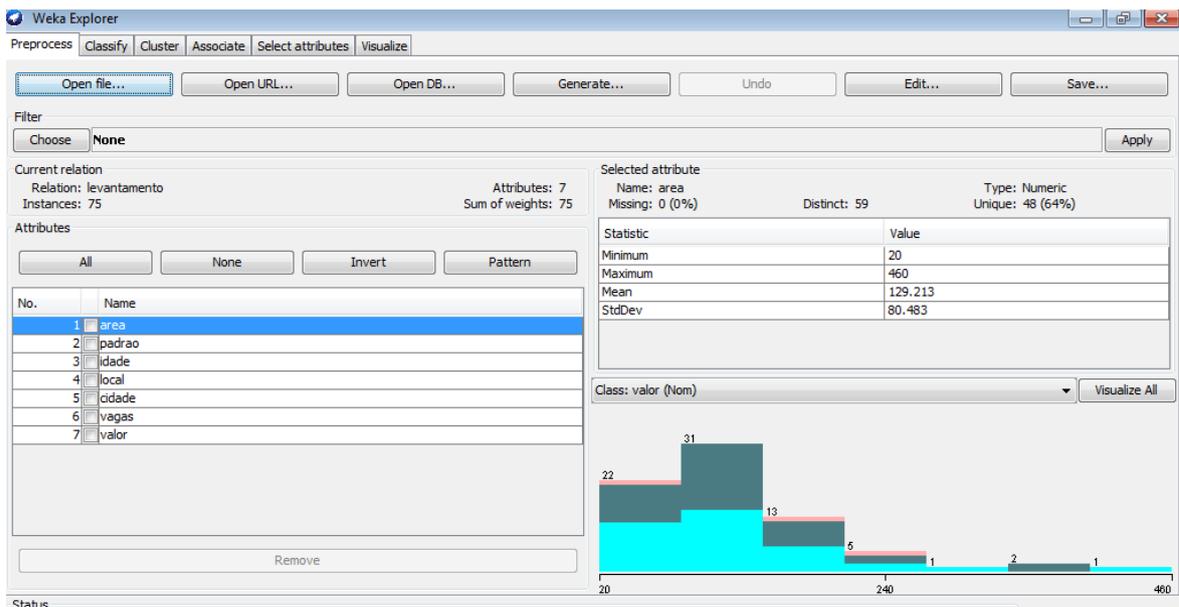


Figura 3.10 – Resultado da leitura do conjunto de dados pelo software WEKA.

Na Figura 3.10 verifica-se que foram processadas 75 instancias, que correspondem aos 75 registros do arquivo de treino/teste, constantes do Apêndice A, contendo 7 atributos, que correspondem ao número de colunas do arquivo, uma para cada atributo de entrada e saída, para a geração da árvore foi utilizado o algoritmo J48.

A Figura 3.10 apresenta outras informações como o maior e o menor valor encontrado no arquivo de entrada para um determinado atributo, que neste caso é a área está assinalada e os seus valores aparecem nas estatísticas: mínimo = 20, máximo = 460 e média = 129,213.

Após o treinamento e teste da base de dados, o WEKA informou os dados da criação da Árvore de Decisão, mostrados na Figura 3.11.

```

=== Summary ===
Correctly Classified Instances      92   %
Incorrectly Classified Instances    8   %

Kappa statistic                    0.8462
Mean absolute error                 0.0508
Root mean squared error             0.1594
Relative absolute error             22.9351 %
Root relative squared error         48.5707 %
Coverage of cases                   100   %
Mean rel. region size (0.95 level)  29.6  %
Total Number of Instances          75

```

Figura 3.11 – Resumo da criação da árvore de decisão pelo software WEKA.

Há muita informação na Figura 3.11, gerada pelo WEKA e definir aquelas mais importantes depende do tipo da aplicação. No caso da geração da árvore de decisão (Classificação), as informações mais relevantes são:

As instâncias classificadas corretamente e incorretamente (Correctly, Incorrectly Classified Instances) mostram a porcentagem de instâncias de teste corretamente e incorretamente classificadas. A porcentagem de instâncias corretamente classificadas indica a precisão nos testes da amostra gerada, que no caso da árvore criada foi de 92%.

A Estatística de Kappa (Kappa statistic) é uma métrica que avalia o nível de concordância de uma tarefa de classificação, através de diferentes técnicas faz-se uma métrica que indica o nível de coesão dos dados classificados. Um valor maior do que 0 significa que os dados possuem uma boa coesão, sendo 1 o valor máximo. Na árvore

gerada foi obtido o valor de 0,8462 para essa métrica, indicando que os valores encontrados possuem uma boa coesão.

As taxas de erro (Erro absoluto médio, Erro quadrático médio da raiz, Erro absoluto relativo e Erro relativo ao quadrado da raiz) são usadas para predição numérica (regressão) em vez de classificação (geração da árvore de decisão). Na predição numérica o erro tem uma magnitude, e essas medidas refletem isso.

O Número total de instâncias (Total Number of Instances) mostra o número total de registros processados, que no caso foram 75.

A Árvore de Decisão criada é mostrada na Figura 3.12, onde observa-se a associação dos ramos formados pelos atributos mais significativos, aqueles com menor entropia, ou seja, com menor aleatoriedade.

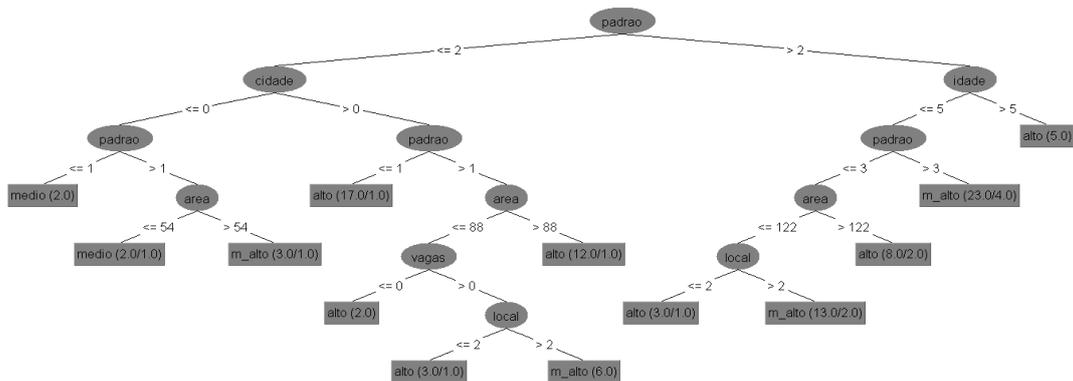


Figura 3.12 – Árvore de decisão criada pelo software WEKA.

A árvore de decisão da Figura 3.12 está impressa em um tamanho maior no Apêndice C.

Ao se percorrer os nós e ramos da árvore de decisão, mostrada na Figura 3.12, verifica-se os padrões de classificação, associando os atributos de entrada do imóvel e a sua classe de saída.

Alguns desses padrões são:

- 1) Se Padrão ≤ 2 e Cidade ≤ 0 e Padrão ≤ 1 Então Valor é Médio;
- 2) Se Padrão ≤ 2 e Cidade ≤ 0 e Padrão > 1 e Área ≤ 54 Então Valor é Médio;
- 3) Se Padrão ≤ 2 e Cidade ≤ 0 e Padrão > 1 e Área > 54 Então Valor é Muito Alto.

À medida que a árvore vai sendo percorrida, as regras de classificação vão sendo definidas. Observa-se que alguns atributos podem aparecer mais de uma vez na mesma regra, neste caso é feita um teste para duas ocorrências diferentes de seus valores.

3.4 - NOVO CONJUNTO DE REGRAS

A partir do percurso dos nós e ramos da árvore gerada, foram criadas **29 regras**, o que tornou viável a implementação do modelo, que anteriormente requeria **810 regras**.

Associando os atributos dessas 29 regras, elas foram formatadas através da estrutura **if-then** (se-então), conforme mostrado na Figura 3.13 e inseridas no modelo do MatLab.

```
1. If (Padrão is inferior) and (Cidade is Ananindeua) then (Valor_m2 is médio) (1)
2. If (Padrão is inferior) and (Cidade is Belém) then (Valor_m2 is alto) (1)
3. If (Área is pequena) and (Padrão is médio) and (Cidade is Ananindeua) then (Valor_m2 is médio) (1)
4. If (Área is média) and (Padrão is médio) and (Cidade is Ananindeua) then (Valor_m2 is m_alto) (1)
5. If (Área is grande) and (Padrão is médio) and (Cidade is Ananindeua) then (Valor_m2 is m_alto) (1)
6. If (Padrão is médio) and (Cidade is Belém) then (Valor_m2 is alto) (1)
7. If (Área is média) and (Padrão is médio) and (Cidade is Belém) then (Valor_m2 is alto) (1)
8. If (Área is grande) and (Padrão is médio) and (Cidade is Belém) then (Valor_m2 is alto) (1)
9. If (Área is pequena) and (Padrão is médio) and (Cidade is Belém) and (Vagas is inferior) then (Valor_m2 is alto) (1)
10. If (Área is pequena) and (Padrão is médio) and (Local is inferior) and (Cidade is Belém) and (Vagas is médio) then (Valor_m2 is alto) (1)
11. If (Área is pequena) and (Padrão is médio) and (Local is medio_inf) and (Cidade is Belém) and (Vagas is médio) then (Valor_m2 is alto) (1)
12. If (Área is pequena) and (Padrão is médio) and (Local is inferior) and (Cidade is Belém) and (Vagas is superior) then (Valor_m2 is alto) (1)
13. If (Área is pequena) and (Padrão is médio) and (Local is medio_inf) and (Cidade is Belém) and (Vagas is superior) then (Valor_m2 is alto) (1)
14. If (Área is pequena) and (Padrão is médio) and (Local is médio) and (Cidade is Belém) and (Vagas is médio) then (Valor_m2 is m_alto) (1)
15. If (Área is pequena) and (Padrão is médio) and (Local is médio_sup) and (Cidade is Belém) and (Vagas is médio) then (Valor_m2 is m_alto) (1)
16. If (Área is pequena) and (Padrão is médio) and (Local is superior) and (Cidade is Belém) and (Vagas is médio) then (Valor_m2 is m_alto) (1)
17. If (Área is pequena) and (Padrão is médio) and (Local is médio) and (Cidade is Belém) and (Vagas is superior) then (Valor_m2 is m_alto) (1)
18. If (Área is pequena) and (Padrão is médio) and (Local is médio_sup) and (Cidade is Belém) and (Vagas is superior) then (Valor_m2 is m_alto) (1)
19. If (Área is pequena) and (Padrão is médio) and (Local is superior) and (Cidade is Belém) and (Vagas is superior) then (Valor_m2 is m_alto) (1)
20. If (Padrão is superior) and (Idade is médio) then (Valor_m2 is alto) (1)
21. If (Padrão is superior) and (Idade is antigo) then (Valor_m2 is alto) (1)
22. If (Padrão is superior) and (Idade is novo) then (Valor_m2 is m_alto) (1)
23. If (Área is média) and (Padrão is médio) and (Idade is novo) then (Valor_m2 is alto) (1)
24. If (Área is grande) and (Padrão is médio) and (Idade is novo) then (Valor_m2 is alto) (1)
25. If (Área is pequena) and (Padrão is médio) and (Idade is novo) and (Local is médio) then (Valor_m2 is m_alto) (1)
26. If (Área is pequena) and (Padrão is médio) and (Idade is novo) and (Local is médio_sup) then (Valor_m2 is m_alto) (1)
27. If (Área is pequena) and (Padrão is médio) and (Idade is novo) and (Local is superior) then (Valor_m2 is m_alto) (1)
28. If (Área is pequena) and (Padrão is médio) and (Idade is novo) and (Local is inferior) then (Valor_m2 is alto) (1)
29. If (Área is pequena) and (Padrão is médio) and (Idade is novo) and (Local is medio_inf) then (Valor_m2 is alto) (1)
```

Figura 3.13 – Regras Fuzzy criadas a partir da árvore de decisão e inseridas no modelo do MatLab.

CAPÍTULO 4

RESULTADOS E DISCUSSÃO

4.1 - APLICANDO OS DADOS

Após a criação do modelo Fuzzy, são nele inferidas as informações de imóveis, tomados aleatoriamente da relação mostrada no Anexo B. Estas informações fazem parte do levantamento de campo realizado pela Faculdade de Engenharia Civil/ITEC/UFGA, mas elas não fizeram parte do conjunto de dados usados para treino e teste da árvore de decisão, mostrados no Anexo A.

Os valores unitários (R\$/m²) calculados pelo modelo são então comparados com os valores observados pela pesquisa de mercado.

Esta inferência e comparação ocorrem em duas etapas, são processados os dados relativos a dois grupos de 12 imóveis cada, separadamente e depois é feita uma comparação do desempenho do modelo Fuzzy, em cada etapa. Estes grupos de imóveis fazem parte do levantamento mostrado no Anexo B.

4.1.1 - Processamento do primeiro grupo de dados

O primeiro grupo de dados submetido ao modelo Fuzzy proposto e seus resultados, está mostrado na Tabela 4.1:

Tabela 4.1 – Primeiro grupo de dados submetido ao modelo Fuzzy e seus resultados.

Item	Área Privativa	Padrão	Idade	Localização	Cidade	Numero Vagas	Valor Calculado no Modelo Fuzzy (R\$/m ²)	Valor Observado no Mercado (R\$/m ²)
18	84,17	2	10	3	1	1	5.146	5.346
26	84,25	3	1	3	1	2	5.146	5.234
22	84,22	1	1	2	0	1	2.667	2.731
13	67,00	1	3	2	1	1	3.902	3.731
37	172,05	4	1	5	1	1	5.630	5.202
82	57,00	1	1	1	1	1	3.902	3.421
89	140,43	3	15	2	1	1	3.902	3.418
96	49,00	2	1	1	0	1	3.067	2.551
1	140,45	5	5	4	1	2	5.962	5.340
54	278,98	1	15	2	1	2	3.902	3.083
94	26,00	2	2	5	1	0	5.146	3.769
92	102,99	2	1	1	1	1	3.124	1.942

A Tabela 4.1 apresenta apenas os atributos selecionados como variáveis linguísticas de entrada (inputs) do modelo, juntamente com os valores calculado e observado. Para consultar outros atributos dos imóveis deve ser utilizada a informação constante na coluna Item, que faz referência à mesma coluna constante no levantamento do Anexo A.

Após inferir os dados de entrada da Tabela 4.1 ao modelo de avaliação proposto, os valores unitários dos imóveis foram calculados e mostrados na mesma tabela, onde os valores calculados no modelo Fuzzy e os observados no mercado são comparados.

As diferenças Absoluta e Relativa entre os valores Calculados e Observados, são mostradas na Tabela 4.2.

Tabela 4.2 – Diferenças entre os valores calculados e observados do grupo 1.

Item	Valor Calculado no Modelo Fuzzy (R\$/m ²)	Valor Observado no Mercado (R\$/m ²)	Diferença	Diferença ²
18	5.146	5.346	-200,32	40.128,10
26	5.146	5.234	-88,42	7.818,10
22	2.667	2.731	-63,94	4.088,32
13	3.902	3.731	170,66	29.124,84
37	5.630	5.202	428,02	183.201,12
82	3.902	3.421	480,95	231.312,90
89	3.902	3.418	483,93	234.188,24
96	3.067	2.551	515,98	266.235,36
1	5.962	5.340	622,02	386.908,88
54	3.902	3.083	819,34	671.318,04
94	5.146	3.769	1.376,77	1.895.495,63
92	3.124	1.942	1.182,06	1.397.265,84
Somatória	51.496	45.769	Somatória	5.347.085,38
Somatória Média	4.291	3.814	Somatória Média	445.590,45
Desvio %	15,56%	17,50%	Desvio Padrão	667,53

Os dados da Tabela 4.2 mostram que o Desvio Padrão da diferença entre os valores calculado e observado foi de 667,53 e a Média dos valores calculados pelo modelo foi de R\$ 4.291,00 e observados no mercado foi de R\$ 3.814,00.

Dividindo-se o Desvio Padrão da diferença pelas Média de cada um dos valores, obtêm-se os percentuais de desvios das médias dos valores calculados e levantados em campo, que foram de 15,56% e 17,50%, respectivamente.

Apesar destas estatísticas apresentarem algumas discrepâncias ou variações elevadas, deve-se levar em conta que os valores observados no mercado, muitas vezes

são frutos do sentimento de valor que cada proprietário ou corretor tem do seu imóvel, o que pode variar em cada caso, dependendo da escala de valorização de cada um, gerando os chamados *outliers*, que são dados que se diferenciam drasticamente de todos os outros, são pontos fora da curva, que fogem da normalidade e que podem causar anomalias nos resultados obtidos por meio de algoritmos e sistemas de análise (AQUARELA, 2017).

Os valores calculados pelo modelo Fuzzy são baseados na análise geral do mercado feita por especialistas, o que pode diferenciar do sentimento de valor dos proprietários.

Observou-se, entretanto, um comportamento padrão do modelo Fuzzy, em relação aos dados de mercado: apesar dos valores calculados apresentarem diferenças para os valores de mercado a **tendência** dos valores calculados segue o padrão apresentado pelo mercado.

Isto pode ser observado na Figura 4.1, onde é traçado um gráfico comparativo dos valores calculados pelo modelo Fuzzy e os levantados no mercado.

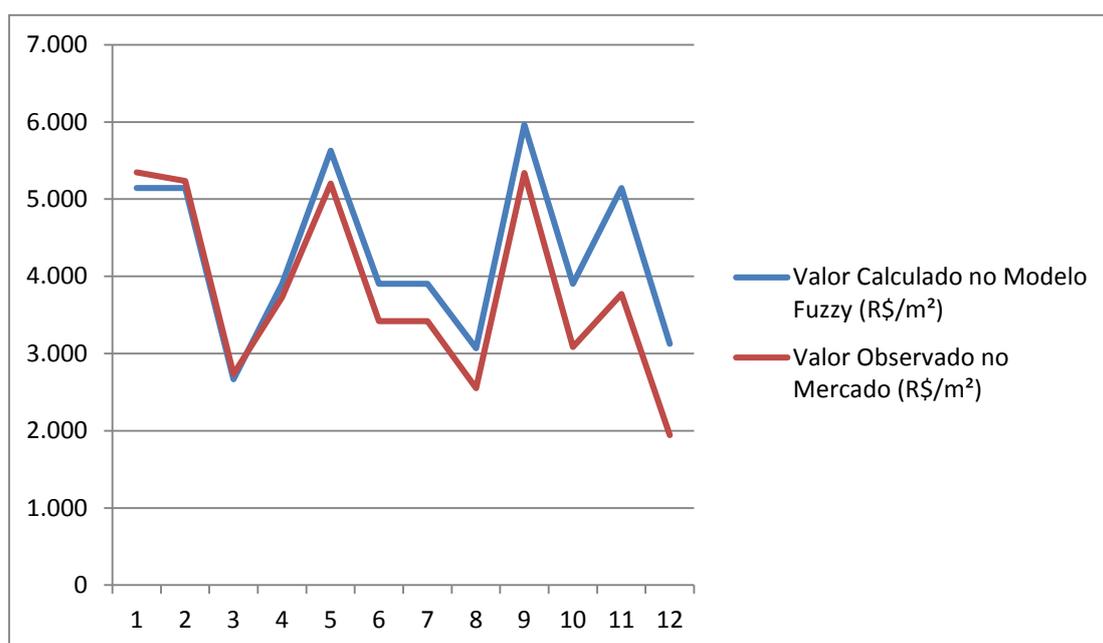


Figura 4.1 – Comparação dos valores calculados e observados no mercado do grupo 1.

A Figura 4.1 mostra o comportamento do valor calculado, que quando não é próximo ao valor observado é proporcional, seguindo a mesma tendência, não apresentando valores extremamente diferentes.

4.1.2 - Processamento do segundo grupo de dados

O segundo grupo de dados submetido ao modelo Fuzzy proposto está mostrado na Tabela 4.3.

Tabela 4.3 – Segundo grupo de dados submetido ao modelo Fuzzy.

Item	Área Privativa	Padrão	Idade	Localização	Cidade	Numero Vagas	Valor Calculado no Modelo Fuzzy (R\$/m ²)	Valor Observado no Mercado (R\$/m ²)
84	84,22	1	1	2	0	1	2.237	2.731
56	191,22	4	15	5	1	1	3.856	3.922
11	47,63	2	5	2	0	1	3.576	3.464
36	65,44	1	2	2	1	1	3.980	3.820
60	135,06	1	35	5	1	1	4.247	3.628
70	171,00	4	1	5	1	3	6.478	5.848
79	51,50	1	1	1	1	1	4.254	3.592
40	86,08	1	10	3	1	1	4.730	4.066
27	156,31	4	1	5	1	2	5.847	5.118
10	85,73	3	2	3	1	2	5.740	4.992
98	22,02	2	5	1	1	0	4.692	3.315
46	81,24	2	3	2	1	2	4.192	3.447

Após submeter os dados da Tabela 4.3 ao modelo de avaliação proposto, os valores do metro quadrado dos imóveis foram calculados e estão mostrados na mesma tabela, onde os valores calculados no modelo Fuzzy e os observados no mercado são comparados.

As diferenças Absoluta e Relativa entre os valores Calculados e Observados, são mostradas na Tabela 4.4.

Tabela 4.4 – Diferenças entre os valores calculados e observados do grupo 2.

Item	Valor Calculado no Modelo Fuzzy (R\$/m ²)	Valor Observado no Mercado (R\$/m ²)	Diferença	Diferença ²
84	2.237	2.731	-493,94	243.976,72
56	3.856	3.922	-66,18	4.379,79
11	3.576	3.464	111,80	12.499,24
36	3.980	3.820	159,71	25.507,28
60	4.247	3.628	618,98	383.136,24
70	6.478	5.848	630,05	396.963,00
79	4.254	3.592	661,77	437.939,53
40	4.730	4.066	664,01	440.909,28
27	5.847	5.118	728,97	531.397,26
10	5.740	4.992	747,58	558.875,86
98	4.692	3.315	1.376,83	1.895.660,85
46	4.192	3.447	745,42	555.650,98
Somatória	53.829	47.944	Somatória	5.486.896,04
Somatória Média	4.486	3.995	Somatória Média	457.241,34
Desvio %	15,07%	16,92%	Desvio Padrão	676,20

Neste novo grupo de dados, da Tabela 4.4, tem-se que o Desvio Padrão da diferença entre os valores calculado e observado foi de 676,20 e a Média dos valores calculados pelo modelo foi de R\$ 4.486,00 e observados no mercado foi de R\$ 3.995,00.

Dividindo-se o Desvio Padrão da diferença pelas Média de cada um dos valores, obtêm-se os percentuais de desvios das médias dos valores calculados e levantados em campo, que foram de 15,07% e 16,92%, respectivamente.

Cabem aqui as mesmas observações feitas para o primeiro grupo de dados, quanto aos *outliers*.

Observou-se, novamente, um comportamento padrão do modelo Fuzzy, em relação aos dados de mercado: apesar dos valores calculados apresentarem diferenças para os valores de mercado a **tendência** dos valores calculados segue o padrão apresentado pelo mercado.

Isto pode ser observado na Figura 4.2, onde é traçado um gráfico comparativo dos valores calculados pelo modelo Fuzzy e os levantados no mercado.

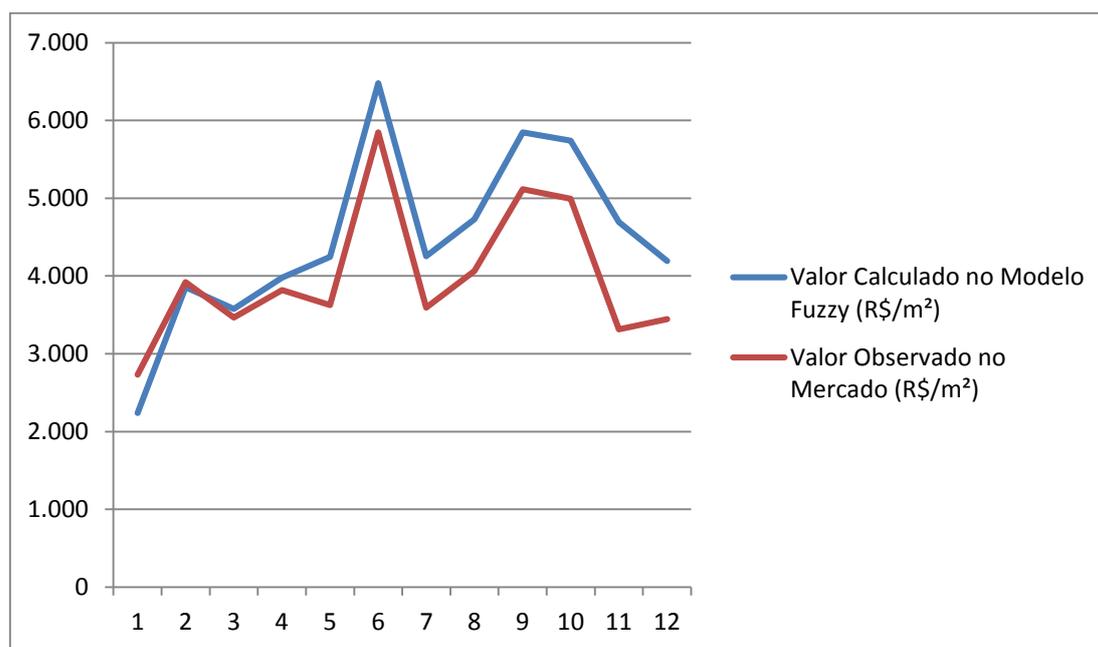


Figura 4.2 – Comparação dos valores calculados e observados no mercado do grupo 2.

A Figura 4.2 mostra o comportamento do valor calculado, novamente com as mesmas características do primeiro grupo de dados: quando não é próximo ao valor observado é proporcional, seguindo a mesma tendência, não apresentando valores extremamente diferentes.

4.2 – DISCUSSÃO DOS RESULTADOS

Observa-se que o comportamento do modelo foi semelhante, no processamento de ambos os grupos.

Apesar dos Desvios Padrão e dos Percentuais de Desvios das médias terem apresentado valores altos, os resultados ocorreram na mesma faixa, com valores próximos: R\$ 667,53, 15,56% e 17,50% para o primeiro grupo e R\$ 676,29, 15,07% e 16,92% para o segundo.

Estes resultados indicam que o modelo Fuzzy proposto apresentou um comportamento uniforme no processamento dos dados dos dois grupos, mantendo os Desvios Padrão e Percentuais de Desvio dentro de uma faixa semelhante.

Este comportamento uniforme, em ambos os casos, é ratificado através das Figuras 4.1 e 4.2, onde o modelo acompanha os valores de mercado, mesmo com diferenças maiores e menores, acompanhando a tendência do mercado, sem apresentar valores extremamente divergentes, o que valida o modelo Fuzzy.

CAPÍTULO 5

CONCLUSÕES E SUGESTÕES

5.1 – CONCLUSÕES

Este trabalho teve como objetivo propor um modelo de avaliação de imóveis, que incorporasse as imprecisões e subjetividades inerentes às variáveis que compõem o mercado imobiliário, para ser utilizado por profissionais que atuam na área. O modelo desenvolvido utilizou a lógica Fuzzy para definir as características subjetivas dos imóveis na forma de variáveis linguísticas, atribuindo a cada uma delas funções de pertinência, refletindo a forma de pensar dos especialistas humanos em avaliações imobiliárias.

Para que o modelo proposto não se tornasse mais um trabalho teórico, na área de avaliação imobiliária utilizando lógica Fuzzy e depois fosse arquivado, devido à impossibilidade de sua implementação prática, causada pela grande quantidade de regras Fuzzy, originalmente geradas, esta pesquisa abordou o uso de técnicas de extração de regras de classificação de atributos relevantes, chamada de *Árvore de Decisão*, para a redução das oitocentas e dez regras, originalmente estabelecidas, para apenas vinte e nove, sem perder o foco na lógica do processo. Esta redução viabiliza o uso prático do modelo proposto, que possui uma arquitetura de dados composta por seis atributos relevantes de entrada (*Área, Padrão, Idade, Local, Cidade e Vagas*), um atributo classe de saída (*Valor do m²*) e vinte e nove elementos na Base de Regras Fuzzy.

Salienta-se o uso, durante as várias etapas deste trabalho, das informações constantes no levantamento de campo realizado pela Faculdade de Engenharia Civil/ITEC/UFPA, obtendo os valores dos 99 imóveis dos municípios de Belém e Ananindeua, diretamente do mercado, conferindo às comparações e simulações feitas no modelo proposto, credibilidade prática.

A pesquisa apresentada neste trabalho abordou uma análise comparativa na utilização do modelo Fuzzy proposto. Foram submetidas informações reais, em duas etapas distintas, gerando valores que foram comparados com dados reais de mercado, e tiveram as suas diferenças levantadas. Na análise dessas diferenças, observou-se que o modelo apresentou um comportamento constante, uniforme em ambas as etapas, resultando nos Desvios Padrão de ambas na ordem de R\$ 600,00 e os Percentuais de Desvio das Médias dos valores calculados e observados na faixa de 15% e 17%,

respectivamente. Os valores do m^2 calculados também seguiram as mesmas tendências dos valores de mercado, sem apresentarem variações díspares. Este comportamento sinaliza positivamente, também, para o uso prático do modelo proposto.

5.2 – SUGESTÕES

Considerando que este estudo limitou-se ao segmento de imóveis residenciais (apartamentos), propõe-se a continuidade das técnicas e metodologias apresentadas, e o estudo da aplicabilidade do modelo proposto, nos seguintes termos:

- a) Inclusão de novas regras criadas por especialistas, além daquelas geradas pela Árvore de Decisão, no estudo atual só foram consideradas as regras geradas pela árvore;
- b) Aumento do número de imóveis utilizados para a comparação dos resultados do modelo, pois o presente estudo limitou-se a dois grupos de doze imóveis;
- c) Expansão do modelo Fuzzy proposto para avaliação de imóveis comerciais e rurais, o estudo atual foi desenvolvido somente para avaliação de imóveis residenciais;
- d) Utilizar uma base de dados maior para validação do modelo, com a redução do Desvio Padrão obtido neste trabalho;
- e) Criação de um aplicativo computacional, utilizando o modelo proposto, visando a utilização prática dos conceitos aqui pesquisados.

REFERÊNCIAS BIBLIOGRÁFICAS

ABNT (Associação Brasileira de Normas Técnicas). NBR 14.653-1: **Avaliação de bens –Parte 1: Procedimentos Gerais**. Rio de Janeiro: ABNT, 2011.

AQUARELA. <https://aquare.la/o-que-sao-outliers-e-como-trata-los-em-uma-analise-de-dados/> . Acesso em 01/12/2017.

ARTERO, Almir O., **Inteligência Artificial Teórica e Prática**, 1ª. Ed., São Paulo, Editora Livraria da Física, 2009.

BASGALUPP, M. P. **LEGAL-Tree: Um algoritmo genético multi-objetivo lexicográfico para indução de árvores de decisão**. Tese de Doutorado - ICMC/USP, São Carlos, 2010.

BOLINELLI, Hermas L., *et al.* **Avaliação de Imóveis Urbanos: Utilização da Lógica Fuzzy para a Consideração de Variáveis Ambientais Singulares**. Belo Horizonte, 2015.

CRUZ, L. F. **Sistematizações da teoria Fuzzy**. Dissertação de Mestrado - Universidade Estadual Paulista. Rio Claro, 1996.

FEITOSA, H. A. **Princípios fundamentais da teoria Fuzzy**. Dissertação de Mestrado - Instituto de Geociências e Ciências Exatas, Universidade Estadual Paulista. Rio Claro, 1992.

DUARTE, André M., **Modelo Geral de Valores Isento de Subjetividade – Caso de Apartamentos na Cidade de Belém**. Dissertação de Mestrado – Universidade Federal do Rio Grande do Sul, 1999.

DUARTE, André M. *et al.* **Um Novo Enfoque na Engenharia de Avaliações: O método do Entendimento Racional**. Belém, 2006.

DURKIN J. **Expert systems: design and development**. New Jersey: Prentice Hall, 1994.

G1 PARÁ. Atlas do IDHM. <http://g1.globo.com/para/>. Acesso em 09/12/2016.

GESTWICKI P. *ID3: History, Implementation, and Applications* - 1997. Disponível em: <http://www.fredonia.edu/students/nixo1903/>. Acesso em: 31/05/2016.

GONZÁLEZ, M. A. S.; FORMOSO, C.T. **Análise Conceitual das dificuldades na Determinação de Modelos de Formação de Preços através de Análise de Regressão**. Universidade Federal do Rio Grande do Sul. Núcleo Orientado para a Inovação da Edificação (NORIE), 2000.

GONZÁLEZ, M. A. S. **Aplicação de Técnicas de Descobrimto de Conhecimento em Bases de Dados e de Inteligência Artificial em Avaliação de Imóveis**. Tese Doutorado em Engenharia Civil –Programa de Pós-Graduação em Engenharia Civil, UFRGS, Porto Alegre -RS. 2002.

HOLSHEIMER M., SIEBES A. **Data Mining: The Search for Knowledge in Databases**. Amsterdam, The Netherlands, 1991. Disponível em: <http://citeseer.nj.nec.com/holsheimer91data.html>. Acesso em 25/09/2016.

IVANQUI, J. **Esteira eletrônica com velocidade controlada por lógica Fuzzy**. Dissertação de Mestrado. Universidade Federal do Paraná, 2005.

JOHNSON, Richard A.; WICHERN, D. W. **Applied Multivariate Statistical Analysis**. New Jersey: Prentice-Hall, 2002. 5ª Ed.

MAIA, Francisco, <http://www.precisao.eng.br/artigos/historia.html>, acesso em 15/12/2015.

MANUAL DE PERÍCIAS, <http://www.manualdepericias.com.br>, acesso em 19/10/2017.

MARÇAL, Rui Francisco Martins, SUSIN, Altamiro A. **O emprego de inteligência artificial como ferramenta de apoio à tomada de decisão na manutenção industrial**. In: III Seminário Catarinense – ABRAMAN 2005. Joinville/SC.

MARRO, Alessandro A.; PELLIS, Antonio N., *et al.* **Lógica Fuzzy: Conceitos e aplicações**. Departamento de Informática e Matemática Aplicada (DIMAp) Universidade Federal do Rio Grande do Norte. Natal, 2010.

MENDONÇA, M.C.; SOLLERO, M.; CURI, E.; AGUIAR, J.B.; QUEIROGA, H.S.; MAIA, E. A.; **Fundamentos de avaliações patrimoniais e perícias de engenharia**. São Paulo, PINI, 1998.

MELLO, Renato, Rubens S. **O Uso de Lógica Fuzzy para Avaliação de Imóveis e Exemplos de Aplicação.** Revista Produção Online, Florianópolis, 2012.

MICHIE D., SPIEGELHALTER D. J., TAYLOR C. C. *Machine learning, Neural and Statistical Classification.* Ellis Horwood, 1994.

OLARU, Cristina, Louis W. **A Complete Fuzzy Decision Tree Technique,** University of Liège, Bélgica, 2003.

OLIVEIRA A. F. N. **Uma metodologia de uso de técnicas de indução para criação de regras de sistemas especialistas** – Programa de Pós-Graduação em Engenharia de Produção. Universidade Federal de Santa Catarina. Florianópolis, Dissertação de Mestrado, 2001.

PELLI, Antonio N. *et al.* **Avaliação de Imóveis Urbanos: Utilização da Lógica Fuzzy e Comparação dos Resultados com o DEA – Envoltória de Dupla Face.** Belo Horizonte, 2015.

POZO A. T. R. **Árvores de Decisão.** Universidade Federal do Paraná, Departamento de Informática - Centro Politécnico, 2002. Disponível em <http://www.inf.ufpr.br/~aurora/>. Acesso em 01/10/2017.

QUINLAN, J. R. **Induction of Decision Trees.** In: **Machine Learning,** 1986.

QUINLAN, J. R. **C4.5: Programs for machine learning,** 1993.

RUSSELL, S. J.; NORVIG, P. **Artificial Intelligence: A modern approach.** PrenticeHall, 1995.

SEGRE A. **Book Review: C4.5: Programs for Machine Learning.** Morgan Kaufmann Publishers, Inc., 1993.

SHAW, Ian S. **Controle e Modelagem Fuzzy.** Edgar Blücher Ltda./FAPESP, São Paulo, 1999.

SILVA, L. M. O. **Uma Aplicação de Árvores de Decisão, Redes Neurais e KNN para a Identificação de Modelos ARMA Não-Sazonais e Sazonais.** Tese de Doutorado, PUC-RIO, 2005.

VIEIRA, R. S. **Protótipo de um sistema de monitoramento remoto inteligente.** Dissertação de mestrado, UFSC, 1999.

WANG, Li-Xin, **A Course in Fuzzy and Systems Control.** New Jersey: Prentice-Hall International, 1997.

WEISS S. M., Kulikowski C. A. **Computer systems that learn: Classification and prediction methods from statistics, neural nets, machine learning, and expert systems.** Morgan Kaufmann Publishers, Inc. 1991.

ZADEH, L. A. **Fuzzy sets and applications.** USA: John Wiley & Sons, 1987.

APÊNDICE A

CONJUNTO DE TREINO/TESTE UTILIZADO PARA A GERAÇÃO DA ÁRVORE DE DECISÃO

Tabela A.1 – Dados utilizados para a geração da árvore de decisão.

Área	Padrão	Idade	Local	Cidade	Vagas	Valor
20	3	3	3	1	0	mm_alto
30	1	35	1	1	0	alto
40	1	30	5	1	0	alto
43	1	25	4	1	0	alto
44	3	1	3	1	1	m_alto
54	2	2	3	1	1	m_alto
56	3	2	3	1	1	m_alto
59	2	1	2	0	2	m_alto
60	2	1	3	1	1	m_alto
63	1	1	1	1	1	alto
65	1	5	2	1	1	alto
67	1	1	3	1	1	alto
69	3	3	2	1	1	alto
72	3	5	5	1	1	m_alto
72	2	1	2	1	2	alto
73	2	1	2	0	1	alto
74	1	1	1	1	1	alto
75	1	1	1	1	2	alto
78	2	1	4	1	1	m_alto
78	2	1	3	1	1	m_alto
79	2	1	2	0	1	m_alto
81	1	20	1	1	1	alto
84	3	1	3	1	2	alto
84	3	1	3	1	2	m_alto
86	2	2	2	1	2	m_alto
87	1	20	3	1	1	alto
88	2	4	4	1	2	m_alto
89	2	1	2	1	2	alto
89	2	2	2	1	2	alto
89	2	2	2	1	1	alto
92	4	3	3	1	2	m_alto
92	3	1	5	1	1	m_alto
92	2	2	4	1	2	alto
93	3	1	5	1	2	m_alto
96	2	10	3	1	1	alto

Tabela A.1 – Continuação.

Área	Padrão	Idade	Local	Cidade	Vagas	Valor
97	2	3	3	1	1	alto
105	3	5	3	1	2	m_alto
107	3	3	2	1	2	alto
107	3	2	2	1	2	m_alto
109	2	5	3	1	1	alto
110	4	1	3	1	2	m_alto
113	4	1	3	1	2	m_alto
118	2	2	1	1	2	alto
118	4	2	4	1	2	m_alto
122	2	20	2	1	1	alto
122	3	1	4	1	2	m_alto
122	3	1	3	1	1	m_alto
127	3	2	3	1	1	alto
137	3	1	3	1	2	alto
137	3	1	3	1	2	alto
137	3	1	3	1	2	m_alto
139	5	1	4	1	2	m_alto
145	4	1	4	1	2	m_alto
147	3	6	4	1	1	alto
149	3	2	4	1	1	alto
154	3	4	2	1	2	m_alto
166	3	1	5	1	3	alto
170	3	3	4	1	2	alto
178	4	4	5	1	2	m_alto
182	4	5	3	1	2	m_alto
182	5	1	5	1	3	m_alto
182	2	14	4	1	2	alto
195	4	1	5	1	3	mm_alto
198	4	2	5	1	3	alto
200	5	1	5	1	3	m_alto
205	5	3	5	1	2	m_alto
218	5	4	3	1	3	m_alto
226	5	1	5	1	3	m_alto
250	4	3	5	1	3	alto
250	5	3	4	1	4	mm_alto
270	5	15	5	1	2	alto
311	4	9	4	1	2	alto
354	4	1	5	1	2	m_alto
359	5	1	5	1	4	m_alto
460	2	20	5	1	3	alto

APÊNDICE B

RESULTADOS OBTIDOS PELO MODELO FUZZY PROPOSTO

Tabela B.1 – Resultados obtidos pelo modelo proposto.

PRIMEIRO GRUPO DE DADOS SUBMETIDO AO MODELO

Item	Área Privativa	Padrão	Idade	Localização	Cidade	Numero Vagas	Valor Calculado no Modelo Fuzzy (R\$/m ²)	Valor Observado no Mercado (R\$/m ²)	Diferença	Diferença ²	
18	84,17	2	10	3	1	1	5.146	5.346	-200,32	40.128,10	
26	84,25	3	1	3	1	2	5.146	5.234	-88,42	7.818,10	
22	84,22	1	1	2	0	1	2.667	2.731	-63,94	4.088,32	
13	67,00	1	3	2	1	1	3.902	3.731	170,66	29.124,84	
37	172,05	4	1	5	1	1	5.630	5.202	428,02	183.201,12	
82	57,00	1	1	1	1	1	3.902	3.421	480,95	231.312,90	
89	140,43	3	15	2	1	1	3.902	3.418	483,93	234.188,24	
96	49,00	2	1	1	0	1	3.067	2.551	515,98	266.235,36	
1	140,45	5	5	4	1	2	5.962	5.340	622,02	386.908,88	
54	278,98	1	15	2	1	2	3.902	3.083	819,34	671.318,04	
94	26,00	2	2	5	1	0	5.146	3.769	1.376,77	1.895.495,63	
92	102,99	2	1	1	1	1	3.124	1.942	1.182,06	1.397.265,84	
							Somatória	51.496	45.769	Somatória	5.347.085,38
							Somatória Média	4.291	3.814	Somatória Média	445.590,45
							Desvio %	15,56%	17,50%	Desvio Padrão	667,53

SEGUNDO GRUPO DE DADOS SUBMETIDO AO MODELO

Item	Área Privativa	Padrão	Idade	Localização	Cidade	Numero Vagas	Valor Calculado no Modelo Fuzzy (R\$/m ²)	Valor Observado no Mercado (R\$/m ²)	Diferença	Diferença ²	
84	84,22	1	1	2	0	1	2.237	2.731	-493,94	243.976,72	
56	191,22	4	15	5	1	1	3.856	3.922	-66,18	4.379,79	
11	47,63	2	5	2	0	1	3.576	3.464	111,80	12.499,24	
36	65,44	1	2	2	1	1	3.980	3.820	159,71	25.507,28	
60	135,06	1	35	5	1	1	4.247	3.628	618,98	383.136,24	
70	171,00	4	1	5	1	3	6.478	5.848	630,05	396.963,00	
79	51,50	1	1	1	1	1	4.254	3.592	661,77	437.939,53	
40	86,08	1	10	3	1	1	4.730	4.066	664,01	440.909,28	
27	156,31	4	1	5	1	2	5.847	5.118	728,97	531.397,26	
10	85,73	3	2	3	1	2	5.740	4.992	747,58	558.875,86	
98	22,02	2	5	1	1	0	4.692	3.315	1.376,83	1.895.660,85	
46	81,24	2	3	2	1	2	4.192	3.447	745,42	555.650,98	
							Somatória	53.829	47.944	Somatória	5.486.896,04
							Somatória Média	4.486	3.995	Somatória Média	457.241,34
							Desvio %	15,07%	16,92%	Desvio Padrão	676,20

APÊNDICE C

ÁRVORE DE DECISÃO GERADA PARA SIMPLIFICAR AS REGRAS DO MODELO

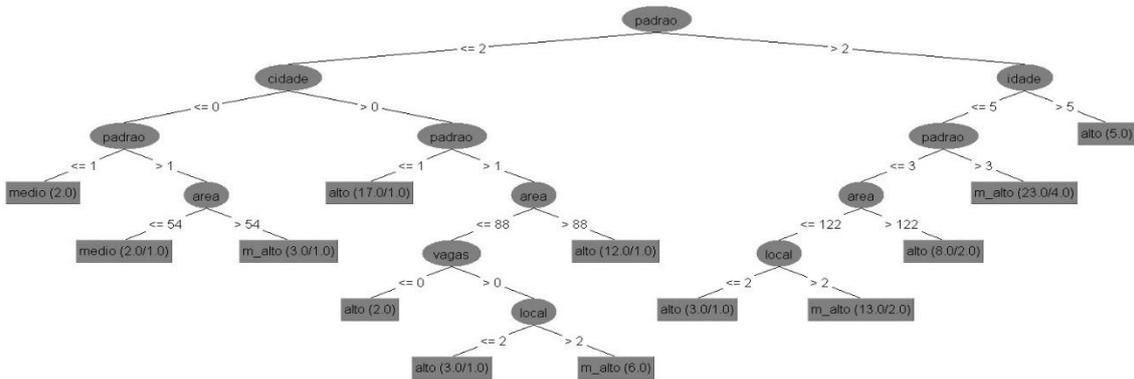


Figura C.1 – Árvore de decisão gerada.

ANEXO I

LEVANTAMENTO DE CAMPO FEITO PELA FACULDADE DE ENGENHARIA CIVIL/ITEC/UFPB

Tabela I.1 – Dados do levantamento de campo.

Item	FOTO/AMOSTRA	ENDEREÇO	EMPRESARIATO	Codificação	BARRIO	CIDADE	CIDADE	DESCRIÇÃO / OBS	Nº Quartos	CONTATO	DATA	DATA	VALOR DE SALGADO	ÁREA PRIVATIVA	Padrão	Área	LOCAL	VALOR m2	Valor Total	Tempo	Região	Observações
2		Rua dos Mundurucu, nº 1551, apto 2002	MARCO	LAT: -7°27'45,87" S LONG: -48°22'25,87" W	Barra-Carmona	Bahia	Bahia	SALA DE ESTAR COM SUGOLA, COZINHA, BANHEIRO, 02 QUARTOS (CIRCULAÇÃO, 3 SUÍTES COM LAVABO, COZINHA, ÁREA DE SERVIÇO E BANHEIRO DE SERVIÇO).	3	97941166 (31) 8429-2122	20/11/15	2	2	138,48	5	3	4	4.594,02	650.000,00	Terminado	3	
3		Ru. Sr. Santa Emília, nº 2514, apto 2002	TOBREVEST	LAT: -7°27'15,11" S LONG: -48°22'31,57" W	Neópolis	Bahia	Bahia	2ª FAV: SALA DE ESTAR, COZINHA, 4 SUÍTES, BANHEIRO COM COZINHEIRO, QUARTO DE SERVIÇO, SALA DE ESTAR, SALA DE JANTAR, LAVABO, COZINHA, ÁREA DE SERVIÇO, QUARTO DE SERVIÇO, BANHEIRO DE SERVIÇO, SALA COM CHURRASQUEIRA, ÁREA DE LANCHES COM PISCINA E COQUELETO.	4	915910 (31) 99811121	20/11/15	1	2	254,40	4	1	5	5.135,44	1.820.000,00	Terminado	1	
4		Ru. Ulisses e Cláudio de Sabotins, nº 1657-1658, Apts 1604	PORTO ALBRE	LAT: -7°28'05,307" S LONG: -48°22'33,57" W	Morro	Bahia	Bahia	SALA DE ESTAR, JANTAR COM SUGOLA, COZINHA, 02 QUARTOS, COZINHEIRO, BANHEIRO SOCIAL, COFA, COZINHA, ÁREA DE SERVIÇO E BANHEIRO DE SERVIÇO.	3	7613498 9394810	20/11/15	1	2	92,51	3	1	5	4.594,24	450.000,00	Terminado	1	
5		Rua Barão do Ceará, 532, apto 2001	URANUS GARIDEN	LAT: -7°27'35,887" S LONG: -48°22'39,887" W	Umarizal	Bahia	Bahia	Sala de estar/jantar com varanda, cozinha, quarto, banheiro social, banheiro social, cozinha, sala de jantar, quarto de serviço, banheiro de serviço e área de serviço.	4	91619 93754551	20/11/15	1	3	200,33	5	1	5	4.738,89	1.350.000,00	Terminado	1	
6		Ru. Engenheiro Fernando Guilherm, nº 2705, Lote 504	SMART GARIDEN BARRA CARMONA	LAT: -7°27'35,887" S LONG: -48°22'45,57" W	Barra-Carmona	Bahia	Bahia	Sala de estar/jantar com varanda, 02 quartos sociais, sendo 01 suíte, 01 banheiro social, 01 banheiro privativo, 01 quarto, 01 banheiro de serviço, 01 quarto de serviço, banheiro de serviço e área de serviço.	3	87666 (31) 82281208	20/11/15	1	2	227,10	3	3	2	4.546,12	497.000,00	Terminado	1	
7		Ru. Nove de Junho, 5615, apto 2002	DUOH	LAT: -7°27'15,11" S LONG: -48°22'31,57" W	Chorão	Bahia	Bahia	03 quartos sociais, 01 banheiro social, 01 banheiro privativo, 01 suíte, 01 banheiro de serviço, 01 quarto de serviço, sala de estar/jantar, cozinha, quarto de serviço e área de serviço.	3	FABRIZO VALINTE (31) 9349-2295	20/11/15	2	2	170,45	3	3	4	4.579,41	475.000,00	Terminado	1	
8		Ru. Angaturama, nº 2952, Bloco B, apto 2002	RIO MENDES	LAT: -7°27'35,887" S LONG: -48°22'39,887" W	Morro	Bahia	Bahia	Sala de estar/jantar, 01 circulação, 02 suítes, 02 banheiros, cozinha, área de serviço, quarto de serviço e banheiro de serviço.	3	Amélia (31) 8425-9202	20/11/15	1	2	112,47	4	1	3	5.132,13	565.000,00	Terminado	1	
9		Rua dos Mundurucu, nº 1571, apto 1002	TOBRE DE MOURA	LAT: -7°27'35,887" S LONG: -48°22'39,887" W	Chorão	Bahia	Bahia	Sala de estar/jantar, cozinha, geladeira, banheiro, circulação, 02 suítes, sendo uma com closet, 02 banheiros, cozinha, área de serviço, quarto de serviço e banheiro de serviço.	3	Luz Estrelado (31) 8333-9265	20/11/15	1	2	121,84	4	5	3	4.754,21	665.000,00	Terminado	1	
12		Ribeirão Augusto Montenegro, nº 4125, Cond. Lote, Bloco D, apto 102	COZINHAS MONTENEGRO	LAT: -7°27'35,887" S LONG: -48°22'45,57" W	Parque Verde	Bahia	Bahia	Sala de estar/jantar, 01 dormitório, 01 suíte, 01 banheiro social, 01 banheiro de serviço, cozinha, circulação, área de serviço e varanda.	2	Marcos (31) 8885-5057	20/11/15	1	1	89,41	3	3	3	4.422,69	307.000,00	Terminado	1	
14		Travessa Humilde, 825, apto 102	PORTO DE ALBUQUERQUE	LAT: -7°27'15,11" S LONG: -48°22'31,57" W	Paulista	Bahia	Bahia	Sala de estar/jantar, cozinha com churrasco, circulação, sala de estar, sala de jantar, um banheiro social, um quarto social, área de serviço, um depósito e banheiro de serviço. Possui área de lazer com playground com 02 piscinas e 04 r. c.	3	Denise (31) 9347-7125	20/11/15	1	2	110,90	4	1	3	4.545,48	610.000,00	Terminado	1	
15		Travessa Dona Romualdo de Sá, 91024, apto 1301	NEON	LAT: -7°27'45,87" S LONG: -48°22'25,87" W	Umarizal	Bahia	Bahia	Sala de estar/jantar, cozinha, 02 suítes, banho, circulação, cozinha, geladeira, área de serviço, banheiro de serviço e quarto de serviço.	4	Neon (31) 3242-0115 9394-7254	20/11/15	2	2	125,88	5	1	3	4.876,62	1.350.000,00	Terminado	1	

Item	FOTO/ARQUITETA	ENDEREÇO	EMPENHAMENTO	Coordenadas	BARRIO	CIDADE	CIDADE	DESCRIÇÃO / OBS	Nº Quartos	CONTRATO	DATA	DATA	VALOR DE GARANTIA	ÁREA PROJETIVA	Padrão	Mês de entrega	LOCAL	VALOR m²	Valor Total	Região	Região Operadora
18		Augusto Montenegro, 900, nº 02, Unid 202	CAMPUS BELLO RESIDENCE	Lat: -23°32'37.3" S Lon: -47°52'59.4" W	Copacabana	Barbém	Barbém	Sala de estar/jantar, copa-cozinha e área de serviço, 2 quartos sociais, 1 banheiro social, 1 banheiro privativo, 1 banheiro/terrace.	3	Miguel (91) 3333-1000 3604897	06/05	1	2	75,00	1	1	1	4.339,39	325.000,00	Trajese	1
17		Condealbino Furtado, 2810, Torre Serenas, Unid 300	PARC PARADISO	Lat: -23°22'04.4" S Lon: -47°52'59.7" W	Chorobaçu	Barbém	Barbém	1 banheiro/terrace, sala de estar/jantar, 2 suites, 2 banheiros privativos, 1 banheiro, geladeira, cozinha, área de serviço e despensa.	2	Yanessa	06/05	1	2	118,20	4	2	4	5.491,11	600.000,00	Trajese	1
19		Rua Domingos Martins, nº 702, unid 1403	NIMES	Lat: -23°26'41.3" S Lon: -47°52'59.7" W	Urca	Barbém	Barbém	Sala de estar / jantar / cozinha gourmet, lavabo, circulação, 4 suites com revestido, cozinha, área de serviço, quarto de serviço e banheiro de serviço.	4	ELIANE (11) 3399-9339	06/05	1	3	181,96	5	1	5	6.181,68	1.125.000,00	Trajese	1
20		Av. Visconde de Souza Franco, 1256, Unid 403	EL GRECO	Lat: -23°26'36.7" S Lon: -47°52'59.7" W	Barbém	Barbém	Barbém	Sala de estar/ jantar, geladeira, cozinha, área de serviço e despensa, 2 quartos sociais, 1 banheiro social, 2 suites, 1 banheiro de serviço, 1 quarto de empregada.	2	PROJETO DALY (11) 3333-2383	06/05	1	1	71,80	3	5	5	5.080,00	365.000,00	Trajese	1
21		Av. Angaturama, 2154, Unid 1805	PORTO DE DOVER	Lat: -23°26'41.3" S Lon: -47°52'59.7" W	Pedreira	Barbém	Barbém	2 quartos sociais, 1 banheiro social, 2 banheiros privativos, 1 banheiro de serviço, 1 banheiro social, 1 sala de estar/jantar, 1 cozinha e 1 área de serviço.	2	ROSE (11) 3334-1334	06/05	2	1	60,27	2	1	3	5.027,38	303.000,00	Trajese	1
23		Av. Saraceni Cunha, nº 333, Unid 204	PAULA DANIELA	Lat: -23°22'31.5" S Lon: -47°52'52.3" W	Neuquén	Barbém	Barbém	Sala de estar, 2 quartos sociais, 1 banheiro social, 1 banheiro de serviço, cozinha e área de serviço.	2	TRACY (11) 3334-9477	06/05	2	0,00	42,80	1	28	4	4.205,81	180.000,00	Trajese	1
24		Av. Elias Pinheiro, 552, Unid 203	RO LISA	Lat: -23°26'34.3" S Lon: -47°52'59.7" W	Mercê	Barbém	Barbém	Sala de estar, circulação, copa-cozinha e área de serviço, 2 quartos sociais, 1 banheiro social, 1 banheiro privativo, 1 banheiro, 1 banheiro/terrace.	1	Sônia (11) 3333-1331	06/05	2	1	58,52	3	2	2	5.801,79	325.000,00	Trajese	1
25		Av. Manoel Barata, nº 241, unid 201 e Tv. 24 de Março, Unid 700	NEW ALVARO	Lat: -23°22'03.3" S Lon: -47°52'59.7" W	Copacabana	Barbém	Barbém	1 quarto social, 1 banheiro social, cozinha.	1	Colombo Souza (11) 3333-2331	06/05	2	0,00	30,00	1	28	1	3.600,00	108.000,00	Trajese	1
26		Av. Tereza Batista, 1495, Unid 804, 801	PORTO DE SINES	Lat: -23°26'31.3" S Lon: -47°52'59.7" W	Morumbi	Barbém	Barbém	Sala de estar/jantar, copa-cozinha, circulação e área de serviço, 2 quartos sociais, 1 banheiro social, 1 banheiro privativo, 1 banheiro de serviço, 1 banheiro/terrace.	2	Patricia (11) 3334-0413	06/05	2	1	63,00	1	1	3	4.000,00	268.000,00	Trajese	1
29		Av. Barão do Rio Branco, 2500, Unid 1004	UNO TOWER	Lat: -23°26'33.3" S Lon: -47°52'59.7" W	Mercê	Barbém	Barbém	Sala de estar, circulação, geladeira, cozinha e área de serviço, 2 banheiros sociais, 1 banheiro de serviço, 2 quartos de serviço.	2	William (11) 3334-1334	06/05	2	2	91,80	4	3	3	5.443,66	500.000,00	Trajese	1
30		Rod. Augusto Montenegro, 4130, Unid. Chequer, Jardim Unid 304 Unid 401C	OSÁGAMA MONTENEGRO	Lat: -23°22'31.5" S Lon: -47°52'59.7" W	Parque Virávia	Barbém	Barbém	1 quarto social, 1 banheiro social, 2 banheiros privativos, 1 banheiro/terrace, sala de estar/jantar, sala de recepção, sala de serviço e cozinha.	3	ANA BARRETO (11) 3333-1333	06/05	1	2	88,66	3	1	2	4.173,29	370.000,00	Trajese	1
31		Av. Humaitá, 1895, Unid 1304	MARANELLO	Lat: -23°26'34.3" S Lon: -47°52'59.7" W	Mercê	Barbém	Barbém	1 quarto social, 1 banheiro social, 2 banheiros privativos, 1 banheiro de serviço, sala de estar/jantar de serviço, sala de estar/jantar de serviço, sala de estar/jantar de serviço.	2	ROSELO (11) 3333-2333	06/05	2	2	84,25	3	1	3	4.688,43	395.000,00	Trajese	1
32		Rod. Augusto Montenegro, Torre 1, unid 202	VERANO RESIDENCIAL	Lat: -23°26'31.3" S Lon: -47°52'59.7" W	Copacabana	Barbém	Barbém	Sala de estar e jantar, copa-cozinha, lavabo, circulação, 1 quarto social, sala de serviço, quarto de serviço.	3	ANDRESSA (11) 3333-1333	06/05	2	2	71,96	2	1	2	4.095,50	295.000,00	Trajese	1
33		Av. 24 de Março, 720, unid 2001	ATHENA GARDEN	Lat: -23°26'30.7" S Lon: -47°52'59.7" W	Urca	Barbém	Barbém	Sala de estar/ jantar, cozinha gourmet, lavabo, circulação, 4 suites, cozinha, área de serviço, quarto de serviço, banheiro de serviço e área de serviço.	3	ANDRESSA (11) 3333-1333	06/05	2	2	144,55	4	1	4	5.372,78	830.000,00	Trajese	1
34		Av. São Francisco, 651, Unid 1802	NAZARY SANTAGO	Lat: -23°26'34.3" S Lon: -47°52'59.7" W	Batavia Campos	Barbém	Barbém	Sala de estar/ jantar com lavabo, circulação, 2 quartos sociais, 1 banheiro social, 1 banheiro privativo, 1 banheiro de serviço, 1 quarto de empregada, sala de serviço, sala de serviço e área de serviço.	3	OSCARDO DOS REIS (11) 3333-1333	06/05	2	1	121,71	2	20	2	3.409,74	415.000,00	Trajese	1
35		Rod. Humaitá de Selvas, 1131, Unid 802	O BONACCI RESIDENCE	Lat: -23°26'30.4" S Lon: -47°52'59.7" W	Urca	Barbém	Barbém	Sala de estar/jantar, 2 suites, 2 banheiros privativos, 1 banheiro de serviço, 1 banheiro/terrace, sala de estar/jantar, sala de serviço.	2	PAZDO (11) 3333-1333	06/05	2	1	92,19	3	1	5	6.274,41	580.000,00	Trajese	1
36		Av. Padre Estelão, 2564, Unid 3002	PORTO DE GÊNIOVA	Lat: -23°26'34.3" S Lon: -47°52'59.7" W	Batavia Campos	Barbém	Barbém	2 quartos sociais, 1 banheiro social, 1 banheiro de serviço, 1 quarto de empregada, sala de estar/jantar com revestido, cozinha e área de serviço.	3	Fernando (11) 3334-0400	06/05	2	1	126,80	3	2	3	4.328,50	548.000,00	Trajese	1
39		Av. Saraceni Lemos, nº 300, unid 302	TORRE DE FARNÉSIA	Lat: -23°26'30.7" S Lon: -47°52'59.7" W	Urca	Barbém	Barbém	Sala de estar com lavabo, um banheiro social, um lavabo, uma sala de jantar, sala de estar, uma sala de serviço, um quarto de empregada e um banheiro de empregada.	4	Edvaldo Andre (11) 3333-1333 9114-9114	06/05	1	3	192,70	4	2	5	4.101,80	590.000,00	Trajese	1
41		Rua Domingos Martins, nº 1493, unid 1304	APRISODIOTS GARDEN	Lat: -23°26'32.7" S Lon: -47°52'59.7" W	Pedreira	Barbém	Barbém	Sala de estar / jantar, lavabo com chuveirinho, circulação, 2 suites, 2 quartos, 1 banheiro social, cozinha, área de serviço, 1 banheiro, banheiro de serviço e 1 quarto para serviço do tipo.	2	ELIANE F. (11) 3333-1333 3222-7777 3222-3333	06/05	2	2	92,39	2	2	4	4.967,39	422.000,00	Trajese	1
42		Trav. Castelo Branco, nº 1194, unid 903	PLAZA COZANNE	Lat: -23°26'32.7" S Lon: -47°52'59.7" W	Guará	Barbém	Barbém	Sala de estar / jantar, lavabo, circulação, 2 quartos, 2 suites, 1 banheiro social, copa/ cozinha, área de serviço, quarto de serviço e banheiro de serviço.	3	PULLANO (11) 3333-1333 3333-1333	06/05	2	1	96,19	2	10	3	3.838,61	350.000,00	Trajese	1
43		Travessa Humaitá, nº 1896, Pa. Mercurio Residência, unid 1004	MARANELLO RESIDENCE	Lat: -23°26'30.4" S Lon: -47°52'59.7" W	Mercê	Barbém	Barbém	Sala de estar, circulação, geladeira, cozinha, sala de serviço, banheiro social, cozinha, área de serviço, banheiro de serviço e sala de circulação e cozinha, sala e quarto.	3	Regina do (11) 3333-1333 3333-1333	06/05	2	2	84,25	3	1	3	5.234,41	441.000,00	Trajese	1
44		Trav. Mauriti, nº 2382, unid 901	SAN TELMO	Lat: -23°26'33.3" S Lon: -47°52'59.7" W	Mercê	Barbém	Barbém	Sala de estar/jantar com lavabo, circulação, 2 suites, quarto social, área de serviço, quarto de serviço e banheiro de serviço.	3	Marta S. (11) 3334-1333 3333-1333	06/05	3	1	121,49	3	1	3	5.469,85	670.000,00	Trajese	1
45		Rua Blumenfeld de Silva, nº 1277, unid 3010 (Lugar)	ALTOS DO UMBRAL, ED. VER ORO	Lat: -23°26'36.7" S Lon: -47°52'59.7" W	Urca	Barbém	Barbém	21 Apartamentos Sala de estar/ jantar com lavabo gourmet, 2 banheiros sociais com closet, quarto, banheiro social, banheiro, revestido, circulação, hall, lavabo, cozinha, banheiro de serviço e 2 unidades de espaço ao 2º pavimento e 2 dependências e cozinha.	3	Albinoza Dias S. (11) 3333-0444 3333-2333	06/05	3	4	398,00	3	1	5	5.510,56	1.980.000,00	Trajese	1

ANEXO II

IMÓVEIS INSERIDOS NO MODELO

Tabela II.1 - Dados inseridos no modelo.

Nº	FOTO/AMOSTRA	ENDEREÇO	EMPRESAMENTO	Codcedimob	BARRIO	CIDADE	CELSID	DESCRIÇÃO / OBS	Nº Quartos	CONTEÚDO	DATA	DATA	VALOR DE GARAGEM	ÁREA PRIVATIVA	Pavão	Andar	LOCAL	VALOR m2	Valor Total	Região	Região Obstruída (Quadrantes)
1		Rua Roseira de São, 1272, JARDIM	SARJUN	LA1-1701841-1 LON-4874948-1	Unicristal	Saúde	1	1 quarto social, 1 banheiro social, 1 banheiro privativo, 1 banheiro de serviço, 1 varanda fechada, sala de estar/jantar, circulação, cozinha e área de serviço.	3	NOBATO (91) 9439-6395	jun/15	1	2	140,45	5	5	4	5.395,86	750.000,00	Traçado	1
10		Rua dos Carpinhos, nº 2742, apto 1402	MISTRA RESIDENCE SERVICE	LA1-1717841-1 LON-4874948-1	Morro	Saúde	1	Sala de estar, cozinha gourmet, quarto, 2 suítes, garagem, cozinha, área de serviço e área de lazer.	2	Marcos (91) 9889-1214	jun/15	1	2	89,79	3	2	3	4.963,41	428.000,00	Traçado	1
11		Rua Mário Cavali, Unif 72, Edif. Escola Via Nova, 9-82	PT COOPERFO	LA1-1717841-1 LON-4874948-1	CasaNova	Araribá	0	Sala de estar, cozinha e área de serviço, 2 quartos sociais, 1 banheiro social, 1 banheiro privativo, 1 varanda fechada.	2	TRACI ROCHA (91) 9163-0579	jun/15	1	1	47,69	2	5	2	3.466,20	395.000,00	Traçado	1
12		Avenida Tancredo Neves, Bloco 1, apto 503	PORTO DE LUZ	LA1-1717841-1 LON-4874948-1	Mesquita	Saúde	1	Sala de estar/jantar, cozinha, circulação, dois dormitórios, um banheiro social, uma copa-cozinha, uma área de serviço e banheiro de serviço. Possui uma sala de estar/jantar com 2 quartos (01).	2	Fátima (91) 91214-4772	jun/15	1	1	63,00	1	3	2	3.372,34	250.000,00	Traçado	1
13		R. Du Odeis, nº 729, UNIF 901, apto 3	RO TUDO	LA1-1717841-1 LON-4874948-1	Paulista	Saúde	1	Sala de estar / jantar, 1 varanda fechada, 1 quarto social, quarto 1 suíte, 1 banheiro social, 1 banheiro privativo, 1 banheiro de serviço, cozinha e área de serviço.	3	JULIANA CARVALHO (91) 9965-6996	jun/15	1	1	84,17	2	10	3	5.246,21	450.000,00	Traçado	1
22		Rodovia BR-324, km 08/47, 910, apto 303, Bloco C	PIRENE RESIDENCIAL TORRE EQUILIBRO	LA1-1717841-1 LON-4874948-1	Centro	Araribá	0	Sala de estar / jantar com cozinha gourmet, sala com dois quartos, um banheiro social, área de circulação, cozinha, área de serviço, banheiro de serviço e varanda integrada à cozinha, sala de quarto.	3	Luciana Jackson (91) 3333-1610	jun/15	2	1	84,22	1	1	2	2.790,34	250.000,00	Traçado	1
26		Parque Humana, nº 2385, apto 1402	MARAMELO	LA1-1717841-1 LON-4874948-1	Morro	Saúde	1	Sala de estar, circulação, cozinha, quarto, sala com varanda, banheiro social, cozinha, área de serviço, banheiro de serviço e varanda integrada à cozinha, sala de quarto.	3	Reynold Alencar (91) 9334-1144	jun/15	2	2	84,26	3	1	3	5.236,41	441.000,00	Traçado	1
27		Av. Senel Bittencourt, nº 3204, apto 702	TORRE WEST	LA1-1717841-1 LON-4874948-1	Nascer	Saúde	1	Sala de estar / jantar com cozinha gourmet, cozinha, circulação, 1 suíte, 1 banheiro social, 1 banheiro de serviço, quarto de serviço e banheiro de serviço.	3	Edson Soares (91) 9193-3777	jun/15	2	2	156,61	4	1	6	5.110,00	800.000,00	Traçado	1
36		Rua Augusto Montenegro, 9701, UNIF 104	RO DAS PEDRAS RESIDENCE	LA1-1717841-1 LON-4874948-1	Parque Verde	Saúde	1	1 quarto social, 1 banheiro social, 1 banheiro privativo, 1 banheiro de serviço, sala de estar e cozinha / área de serviço.	2	Geovani (91) 91131-3013	jun/15	2	1	65,44	1	2	2	3.823,29	250.000,00	Traçado	1
37		Rua Roseira de São, 1272, UNIF 901, apto 503, Bloco C, Torre 4	ALTOS DO UNIVERSAL	LA1-1717841-1 LON-4874948-1	Unicristal	Saúde	1	1 quarto social, 1 banheiro social, 1 banheiro privativo, 1 banheiro de serviço, 1 quarto de serviço, 1 banheiro social, 1 banheiro privativo, sala de estar/jantar, circulação, cozinha, área de serviço.	4	Roberto (91) 9163-6616	jun/15	2	1	172,08	4	1	5	5.201,88	895.000,00	Traçado	1
40		Avenida Marçal de Carvalho, Nº 85, apto 1106	RO GARDES	LA1-1717841-1 LON-4874948-1	Paulista	Saúde	1	Varanda, sala de estar / jantar, 1 quarto social, copa-cozinha, área de serviço, banheiro social e banheiro de serviço.	3	Paulina T. Silva Costa (91) 3233-1211 3233-1012	jun/15	1	1	80,08	1	10	3	4.865,90	350.000,00	Traçado	1
41		Rua Augusto Montenegro, Nº 101, UNIF, Torre 2, apto 113	PT MARANTE DO PARQUE	LA1-1717841-1 LON-4874948-1	Mesquita	Saúde	1	Sala de estar / jantar, cozinha gourmet, circulação, 1 quarto social, 1 banheiro social, 1 banheiro de serviço, sala de estar e cozinha, cozinha e área de serviço.	3	Marcelo (91) 9121-1214	jun/15	3	2	81,24	2	3	2	3.486,56	280.000,00	Traçado	1
54		Rua Arcobispo Manoel, nº 300, Torres, apto 801	SANEL	LA1-180726381-1 LON-4874948-1	Bela Vista	Saúde	1	Sala de estar / jantar / 1 quarto social, cozinha, 1 banheiro social, sala de estar com closet, sala, cozinha, quarto de serviço, banheiro de serviço e varanda.	3	Daniel (91) 844-9233 8336-7856	ago/15	2	2	276,96	1	15	3	3.063,65	860.000,00	Traçado	1
56		Torreana Dom Romualdo de Sales, nº 1301, apto 101	MALAGA	LA1-180726381-1 LON-4874948-1	Unicristal	Saúde	1	Sala para 2 dormitórios com cozinha, sala de estar, banheiro, 1 quarto social com closet e varanda, sala, cozinha, quarto de serviço, banheiro de serviço e banheiro de serviço.	3	Thales Lobato (91) 3333-3207 3446-3200	ago/15	2	1	181,12	4	15	5	3.922,18	750.000,00	Traçado	1
60		Av. São de Aguiar, nº 816, B1, apto 1001	JARDIM FRANÇA	LA1-1717841-1 LON-4874948-1	Nascer	Saúde	1	Sala de estar, sala de estar / jantar, circulação, 1 quarto social, 1 banheiro social, cozinha, área de serviço, banheiro de serviço e quarto de serviço.	3	Marcos Antonio (91) 8889-2190 3313-1300	ago/15	3	1	135,06	1	25	5	3.628,02	495.000,00	Traçado	1
70		R. Magalhães Barata	TERRAZZO	LA1-1717841-1 LON-4874948-1	São José	Saúde	1	SALA AMPLA COM LAVABO, SALA COM SACADA, HALL SOCIAL E 02 SERVIÇOS. CIRCULADORA, COZINHA, DESPACHA, QUARTO E BANHEIRO DE SERVIÇO, ÁREA DE SERVIÇO.	3	Eda Dornelles (91) 9889-0388	ago/15	3	3	170,00	4	1	5	5.847,81	1.000.000,00	Oferta	0
71		Torre do Teard, s/nº (Rua Augusto Montenegro)	ALESSO MONTENEGRO	LA1-1717841-1 LON-4874948-1	Parque Verde	Saúde	1	Sala de estar/jantar com cozinha, 1 quarto, 1 suíte, 1 banheiro social, cozinha e área de serviço.	2	Cláudia Tereza (91) 3333-8080	ago/15	3	1	51,50	1	1	1	3.953,23	185.000,00	Oferta	0
82		Rodovia Augusto Montenegro	TOTAL LIFE	LA1-1717841-1 LON-4874948-1	Tanque	Saúde	1	HALL, SALA DE ESTAR/JANTAR COM VITRAJADA, 02 QUARTOS, 02 BANHOS, 01 SUÍTE, COZINHA, 01 ÁREA DE SERVIÇO.	3	César (91) 9889-1688	ago/15	3	1	57,00	1	1	1	3.421,05	135.000,00	Oferta	0
84		Rua BR 316, KM 08, apto 402, Bloco 1	PIRENE RESIDENCIAL TORRE ALDEIA	LA1-1717841-1 LON-4874948-1	Centro	Araribá	0	100, sala de estar / jantar com cozinha gourmet, área de circulação, 1 suíte, 02 quartos, 1 banheiro social, cozinha e área de serviço com banheiro.	3	Paulo Roberto (91) 9121-0284	ago/15	3	1	84,22	1	1	1	2.790,34	250.000,00	Traçado	1
88		Torreana de Maracá, nº 370, UNIF 802	PIREMONI	LA1-180726381-1 LON-4874948-1	São José	Saúde	1	1 quarto social, 1 banheiro social, 1 banheiro privativo, 1 banheiro de serviço, sala de estar, sala de circulação, 1 banheiro social, 1 banheiro privativo, sala de estar/jantar com cozinha gourmet, sala de circulação, sala de estar/jantar com cozinha gourmet e área de serviço.	4	Jeff (91) 9889-3402	ago/15	3	1	149,49	3	15	3	3.418,07	495.000,00	Traçado	1
92		Torreana Moura, apto 301, nº 1039	ERLUS	LA1-180726381-1 LON-4874948-1	Paulista	Saúde	1	Sala de estar / jantar, 01 suíte, 02 quartos, banheiro social, cozinha e área de serviço.	3	Paulo Sérgio (91) 9889-3402 8888-4142	ago/15	3	2	102,30	3	1	3	3.441,34	300.000,00	Traçado	1
94		Av. Senel Bittencourt, 85	SOFF INA	LA1-180726381-1 LON-4874948-1	Bela Vista	Saúde	1	01 Quarto e banheiro privativo	1	Viviane Raul (91) 3333-6666	ago/15	3	0	26,00	2	2	5	3.763,23	98.000,00	Oferta	0
96		Av. Itália, Quarteirão nº 220 (Unif. Rocas)	ROQUES VESPALES	LA1-1717841-1 LON-4874948-1	CasaNova	Araribá	0	HALL, Sala de estar/jantar, 02 quartos, banheiro e área de serviço.	2	Cláudia Rodrigues (91) 9889-0388	ago/15	3	1	48,00	2	1	1	2.551,03	125.000,00	Oferta	0
98		Rua 15 de novembro, nº 39 - Duque de Te. 7 de setembro, UNIF 901, Tor 1	NAZARE	LA1-1717841-1 LON-4874948-1	CasaNova	Saúde	1	HALL sala de estar / jantar, 1 quarto, banheiro social e cozinha.	1	Vanessa Soares (91) 9889-1435	ago/15	3	0	22,02	2	5	1	3.515,17	75.000,00	Traçado	1