



UNIVERSIDADE
ESTADUAL DE LONDRINA

ESTEFÂNIA MAYUMI FUZYI

MODELAGEM FUZZY PARA PROBLEMAS DE
CLASSIFICAÇÃO COM AMOSTRAS NÃO-FACTÍVEIS

Londrina
2017

ESTEFÂNIA MAYUMI FUZYI

MODELAGEM FUZZY PARA PROBLEMAS DE
CLASSIFICAÇÃO COM AMOSTRAS NÃO-FACTÍVEIS

Dissertação apresentada ao Programa de
Mestrado em Ciência da Computação da
Universidade Estadual de Londrina para obtenção
do título de Mestre em Ciência da Computação.

Orientador: Prof. Dr. Sylvio Barbon Jr.

Londrina
2017

ESTEFÂNIA MAYUMI FUZYI

**MODELAGEM FUZZY PARA PROBLEMAS DE CLASSIFICAÇÃO
COM AMOSTRAS NÃO-FACTÍVEIS**

Dissertação apresentada ao Programa de Mestrado em Ciência da Computação da Universidade Estadual de Londrina para obtenção do título de Mestre em Ciência da Computação.

BANCA EXAMINADORA

Orientador: Prof. Dr. Sylvio Barbon Jr
Universidade Estadual de Londrina - UEL

Prof. Dr. Lucas Dias Hiera Sampaio
Universidade Tecnológica Federaldo Paraná -
UTFPR

Profa Dra. Priscila Tiemi Maeda Saito
Universidade Tecnológica Federaldo Paraná -
UTFPR

Prof. Dr. Douglas Fernandes Barbin
Universidade Estadual de Campinas –
UNICAMP

Londrina, 14 de fevereiro 2017.

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UEL

Fuzyi, Estefânia Mayumi.

Modelagem Fuzzy para Problemas de Classificação com Amostras Não-Factíveis / Estefânia Mayumi Fuzyi. - Londrina, 2017.
75 f. : il.

Orientador: Sylvio Barbon Jr.

Dissertação (Mestrado em Ciência da Computação) - Universidade Estadual de Londrina, Centro de Ciências Exatas, Programa de Pós-Graduação em Ciência da Computação, 2017.

Inclui bibliografia.

1. Sistemas de Inferência Fuzzy - Tese. 2. Reconhecimento de Padrões - Tese. 3. Aprendizado de Máquina - Tese. 4. Qualidade de Alimentos - Tese. I. Barbon Jr, Sylvio. II. Universidade Estadual de Londrina. Centro de Ciências Exatas. Programa de Pós-Graduação em Ciência da Computação. III. Título.

AGRADECIMENTOS

Agradeço a todos os professores e funcionários do Departamento de Computação, que de alguma forma contribuíram com a minha formação e à Fundação CAPES por financiar meus estudos durante o período do programa.

FUZYI, E. M. **Modelagem Fuzzy para Problemas de Classificação com Amostras Não-Factíveis**. 75 p. Dissertação de Mestrado (Mestrado em Ciência da Computação) – Universidade Estadual de Londrina, Londrina-PR, 2017.

RESUMO

Alguns modelos de classificação não lidam muito bem com amostras reais, devido a rigidez de seus limites ou a existência de amostras que não se encaixam perfeitamente no intervalo esperado. Em contrapartida, a Lógica Fuzzy apresenta vantagens como a transição gradual entre a alta e baixa pertinência da amostra ao conjunto, além de ser um modelo descritivo de fácil entendimento e análise dos resultados. Dessa forma, a proposta desse trabalho é verificar o uso dos modelos Fuzzy para classificação de amostras não-factíveis e comparar seu desempenho com algoritmos de Aprendizado de Máquina. Com esse objetivo, os métodos foram aplicados ao problema de avaliação de qualidade de carne suína, que apresenta diferentes padrões de avaliação e amostras não-factíveis quando utilizada a lógica Clássica para classificação. De acordo com os experimentos, no pior caso a lógica Clássica foi capaz de classificar 5,88% de um dataset de 306 amostras, enquanto a Fuzzy Top-Down subiu para 18,30%. Em relação aos algoritmos de Aprendizado de Máquina, a *Random Forest* (RF) apresentou melhores resultados em comparação aos demais, com acurácias de 100% na maioria dos casos. Porém, o Fuzzy Top-Down foi capaz de igualá-la ou até mesmo superá-la, como quando a RF obteve acurácia de 63,04% e o Fuzzy Top-Down de 100%, o que comprova sua aderência ao problema.

Palavras-chave: Sistemas de Inferência Fuzzy. Reconhecimento de Padrões. Aprendizado de Máquina. Qualidade de Alimentos

FUZYI, E. M. **Fuzzy Modelling for Classification Problems with Infeasible Samples**. 75 p. Master's Thesis (Master in Science in Computer Science) – State University of Londrina, Londrina–PR, 2017.

ABSTRACT

Some classification models does not handle very well with real samples, because of its rigid thresholds and samples that do not fit perfectly in the expected interval. On the other hand, the Fuzzy Logic has advantages like the smooth transition between full membership and no membership to a set, besides being a descriptive model with results of easy understanding. In this way, this paper verifies the use of Fuzzy models to classify infeasible samples and compares its performance with Machine Learning algorithms. For this purpose, theses methods were applied to pork quality assessment, which has different class grades and infeasible samples when classified by Classical Logic. According to the experiments, in the worst case, the Classical Logic was able to classify only 5,88% of a dataset with 306 samples, while Fuzzy Top-Down increased to 18,30%. Regarding Machine Learning algorithms, the Random Forest (RF) showed better results than others, with 100% accuracy in most cases. However, Fuzzy Top-Down achieved the same or better results than RF, for instance when RF got 63,04% and Fuzzy Top-Down 100%, which address its adherence to the problem.

Keywords: Fuzzy Inference Systems. Pattern Recognition. Machine Learning. Food Quality

LISTA DE ILUSTRAÇÕES

Figura 1 – Exemplo de amostras linearmente separáveis e não-linearmente separáveis [1]	24
Figura 2 – Entrada $x = 0,5$ para as diferentes formas de funções de pertinência. (a) Gaussiana com $\sigma = 0,2$ e $c = 0,5$. (b) Trapezoidal com $p_1 = 0,3$, $p_2 = 0,4$, $p_3 = 0,6$ e $p_4 = 0,7$. (c) Triangular com $p_1 = 0,3$, $p_2 = 0,5$ e $p_3 = 0,7$	31
Figura 3 – Função de Pertinência para o Conjunto de Altura de Pessoas	32
Figura 4 – Sistema de Inferência Fuzzy. Adaptado de [2]	33
Figura 5 – Fuzzificação para o pH	37
Figura 6 – Aplicação do Operador Fuzzy AND na Regra	37
Figura 7 – Aplicação do Método de Implicação (MIN) na Função de Pertinência de Saída	38
Figura 8 – Diagrama de Inferência Fuzzy, demonstrando a quantidade de regras criadas considerando o número de classes	40
Figura 9 – Seis graus de maturação do tomate, de acordo com o USDA [3]	41
Figura 10 – Diagrama da Metodologia Proposta	45
Figura 11 – Funções de Pertinência construídas para cada padrão	52
Figura 12 – Exemplo de Avaliação de Desempenho de Múltiplos Modelos, com ran- queamento de Friedman e cálculo da CD com Nemenyi [4]	54
Figura 13 – Acurácia para os Datasets 1 a 6, em relação aos algoritmos de aprendi- zado de máquina (Fuzzy-CHI, Fuzzy-W, J48, k-NN, MLP, SVM, NB e RF) e padrões de avaliação de carne (Faucitano, Joo, Kauffman e Warner)	59
Figura 14 – Acurácia para os Datasets 7 a 10, em relação aos algoritmos de aprendi- zado de máquina (Fuzzy-CHI, Fuzzy-W, J48, k-NN, MLP, SVM, NB e RF) e padrões de avaliação de carne (Faucitano, Joo, Kauffman e Warner)	60
Figura 15 – Análise de Desempenho dos Algoritmos de Aprendizado de Máquina (Fuzzy-CHI, Fuzzy-W, J48, k-NN, MLP, SVM, NB e RF) por Dataset para o Padrão Faucitano, com $CD = 3,32$ e $\alpha = 0,05$	60
Figura 16 – Análise de Desempenho dos Algoritmos de Aprendizado de Máquina (Fuzzy-CHI, Fuzzy-W, J48, k-NN, MLP, SVM, NB e RF) por Dataset para o Padrão Joo, com $CD = 3,32$ e $\alpha = 0,05$	61
Figura 17 – Análise de Desempenho dos Algoritmos de Aprendizado de Máquina (Fuzzy-CHI, Fuzzy-W, J48, k-NN, MLP, SVM, NB e RF) por Dataset para o Padrão Kauffman, com $CD = 3,32$ e $\alpha = 0,05$	61

Figura 18 – Análise de Desempenho dos Algoritmos de Aprendizado de Máquina (Fuzzy-CHI, Fuzzy-W, J48, k-NN, MLP, SVM, NB e RF) por Dataset para o Padrão Warner, com $CD = 3,32$ e $\alpha = 0,05$	61
Figura 19 – Amostras não-factíveis para Faucitano	62
Figura 20 – Amostras não-factíveis para Kauffman	63
Figura 21 – Amostras não-factíveis para Warner	63
Figura 22 – Lógica Clássica vs. Top-Down	64

LISTA DE TABELAS

Tabela 1 – Exemplo de dataset para classificação	24
Tabela 2 – Operadores Fuzzy T-norm e S-norm	34
Tabela 3 – Padrão Warner [5]	36
Tabela 4 – Padrões de Qualidade para Carne suína	47
Tabela 5 – Quantidade total de amostras e medidas de dispersão para os atributos das bases de dados com amostras reais	48
Tabela 6 – Medidas de dispersão para pH, L* e CRA das bases sintéticas geradas para os padrões Faucitano, Joo, Kauffman e Warner	49
Tabela 7 – Algoritmos utilizados e seus pacotes R	50
Tabela 8 – Parâmetros de configuração para a geração do modelo FRBS com a abordagem Top-Down	51
Tabela 9 – Regras IF-THEN para o padrão de Faucitano	52
Tabela 10 – Regras IF-THEN para o padrão de Joo	53
Tabela 11 – Regras IF-THEN para o padrão de Kauffman	53
Tabela 12 – Regras IF-THEN para o padrão de Warner	53
Tabela 13 – Porcentagem de Amostras Classificadas pela Lógica Clássica	56
Tabela 14 – Porcentagem de Amostras Classificadas pela Abordagem Top-Down	57
Tabela 15 – Acurácia da Classificação Abordagem Top-Down (Acc e total de amostras)	57
Tabela 16 – Comparação das acurácias da classificação dos Datasets Originais, entre o Fuzzy Top-Down e os oito algoritmos de aprendizado de máquina	66

LISTA DE ABREVIATURAS E SIGLAS

AM	Aprendizado de Máquina
CD	<i>Critical Difference</i>
CRA	Capacidade de Retenção de água
DFD	<i>Dark, Firm, Dry</i>
DL	<i>Drip Loss</i>
FIS	<i>Fuzzy Inference System</i>
FPW	<i>Filter Paper Wetness</i>
FRBS	<i>Fuzzy Rule-Based System</i>
GLCM	<i>Gray Level Co-Occurrence Matrix</i>
GLGCM	<i>Gray Level-Gradient Co-Occurrence Matrix</i>
IC	Inteligência Computacional
k-NN	<i>k-Nearest Neighbors</i>
LANA	Laboratório de Análise de Alimentos e Nutrição Animal
LDA	<i>Linear Discriminant Analysis</i>
MLP	<i>Multilayer Perceptron</i>
NIRS	<i>Near-infrared Spectroscopy</i>
NB	<i>Naïve Bayes</i>
PFN	<i>Pale, Firm, Non-exudative</i>
PSE	<i>Pale, Soft, Exudative</i>
RF	<i>Random Forest</i>
RFN	<i>Red, Firm, Non-exudative</i>
RBS	<i>Rule-Based System</i>
RNA	Rede Neural Artificial
RSE	<i>Red, Soft, Exudative</i>

USDA *United States Department of Agriculture*

SVM *Support Vector Machine*

SUMÁRIO

1	INTRODUÇÃO	19
1.1	Objetivo	20
1.2	Organização do Trabalho	21
2	FUNDAMENTAÇÃO TEÓRICA	23
2.1	Modelagem Preditiva	23
2.2	Factibilidade	24
2.3	Aprendizado de Máquina	25
2.4	Sistemas Baseados em Regras	28
2.5	Fundamentos Fuzzy	29
2.5.1	Conjuntos Fuzzy	29
2.5.1.1	Funções de Pertinência	30
2.5.1.2	Operadores de Conjuntos	32
2.5.2	Sistemas de Inferência Fuzzy	32
3	PROCESSO DE INFERÊNCIA FUZZY	35
3.1	Passo 1: Fuzzificar as Entradas do FIS	36
3.2	Passo 2: Aplicar o Operador Fuzzy	36
3.3	Passo 3: Aplicar Método de Implicação	38
3.4	Passo 4: Defuzzificação e Agregação	38
4	TRABALHOS RELACIONADOS	41
5	METODOLOGIA	45
5.1	Avaliação de Qualidade de Carne Suína	46
5.2	Base de Dados	46
5.3	R	49
5.4	Modelos	50
5.5	Avaliação de Desempenho de Modelos Multidomínio	53
6	RESULTADOS E DISCUSSÃO	55
6.1	Experimento 1: Lógica Clássica	55
6.2	Experimento 2: Top-Down	56
6.3	Experimento 3: Bottom-Up	58
6.4	Análise dos Graus de Pertinência das Amostras Não-Factíveis	61
6.5	Comparação Lógica Clássica <i>vs.</i> Top-Down	64
6.6	Comparação Top-Down <i>vs.</i> Bottom-Up	65

7	CONCLUSÃO	67
	REFERÊNCIAS	69
	Trabalhos Publicados pelo Autor	75

1 INTRODUÇÃO

Na lógica Clássica, também denominada lógica Aristotélica, uma proposição só pode assumir dois valores verdade, ou são verdadeiras ou são falsas. Esses sistemas são denominados binários ou também lógica bivalente, pois envolvem dois valores lógicos [6]. O aprendizado *bayesiano* e outros modelos de probabilidade também seguem a lógica bivalente, pois é utilizada para expressar a probabilidade de uma proposição se tornar verdadeira, porém cada fato apenas é considerado como verdadeiro ou falso.

Outro tipo de lógica é a denominada polivalente, que inicialmente foi utilizada no Princípio da Incerteza da física quântica. Nesse caso, ela apresenta três valores: verdadeiro, falso e indeterminado. Uma extensão dessa lógica é considerar o valor 0 como falso, 1 como verdade e valores entre 0 e 1 como graus de verdade, como ocorre na lógica Fuzzy.

A principal diferença entre a lógica Clássica e a lógica Fuzzy é que na primeira tem-se a certeza que o valor verdade da proposição será verdadeiro ou falso. Por exemplo, quando um fator tem valor de probabilidade de 0,5 é provável que ele seja tanto verdadeiro quanto falso, mas sempre será verdadeira ou falsa, nunca ambos ou nenhum. Enquanto na lógica Fuzzy por exemplo, esse seria o grau em que essa afirmação é verdadeira, podendo ser nem verdadeiro nem falso ou tanto verdadeiro quanto falso.

A teoria de lógica Fuzzy, fundada por Zadeh [7] tem sido amplamente utilizada como uma ferramenta poderosa para modelagem do pensamento humano. Pois, da mesma forma que as funções cognitivas do cérebro são baseadas em graus relativos das informações adquiridas pelos sensores biológicos, a lógica Fuzzy lida com o conceito linguístico de função de pertinência, ao contrário das proposições bivalentes [8]. Os princípios da lógica Fuzzy podem ser aplicados tanto para o controle de processos físicos e químicos, como a temperatura e a corrente elétrica, quanto para reconhecimento de padrões. Nesse último caso, pode ser utilizado o Sistema de Inferência Fuzzy (*Fuzzy Inference System* - FIS) que trabalha com regras SE-ENTÃO (*IF-THEN*).

FIS para classificação e predição são utilizados em diversas áreas do conhecimento, como arquiteturas de predição de violação de contratos virtuais [9], classificação da parte inferior do corpo feminino, baseado em medidas antropométricas, para auxiliar na indústria do vestuário na China [10], avaliação de qualidade do serviço público hospitalar da região da Sicília, na Itália [11]. Além dos alimentos, tal como avaliação do nível de risco de segurança alimentar [12], determinação do tempo de descarte do óleo de girassol usado para fritura, junto com *e-nose* [13] e a classificação de mangas, maçãs, tâmaras e tomates [14, 15, 16, 17, 3].

Recentemente, a importância da qualidade da carne tem aumentado, tanto quanto

a quantidade produzida [18], devido ao crescimento da demanda do consumidor por carnes mais saudáveis [19]. No entanto, a indústria de carne não tem uma única definição para qualidade, devido às variações de mercado [20]. Além disso, a qualidade da carne suína é altamente variável e possui diversas metodologias de classificação. Além da avaliação laboratorial ser custosa, consumir tempo e depender de pessoas treinadas e avaliação subjetiva. Sendo assim, métodos alternativos e análises não-destrutivas do alimento com o uso de ferramentas computacionais [21].

O uso de Inteligência Computacional (IC) para classificação de qualidade de alimentos tem sido amplamente discutido [22, 23, 24, 25]. A principal vantagem da IC é a facilidade de avaliação no ambiente industrial, ser mais rápido que o ser humano, mais preciso e ter custos menores [26]. Entre os diversos algoritmos de Aprendizado de Máquina (AM) de IC para avaliação de qualidade de alimentos, alguns se destacam, como algoritmos de Clusterização, Sistemas Fuzzy Baseado em Regras (*Fuzzy Rule-Based Systems* - FRBS), *k-Nearest Neighbors* (k-NN), Máquinas de Vetor de Suporte (*Support Vector Machine* - SVM) e a Perceptron Multicamadas (*Multilayer Perceptron* - MLP) [27].

Técnicas de AM têm sido aplicadas em diversos tipos de pesquisa de qualidade de alimentos, principalmente em tarefas de predição e classificação. Combinadas ou não, elas têm sido amplamente utilizadas para verificar deterioração, defeitos, fraudes ou predizer os parâmetros de qualidade mais importantes para avaliação da carne [28, 29, 30, 31].

Os métodos de avaliação de carne suína quanto à qualidade, são estabelecidos de acordo com a teoria de conjuntos tradicional ou *crisp*. Eles apresentam limites rígidos entre os intervalos dos atributos, o que torna a transição entre as classes abrupta e dificulta a classificação de amostras que não se ajustam perfeitamente aos intervalos esperados. Ao contrário dos conjuntos fuzzy, que apresentam uma transição mais gradual entre a completa pertinência e a não pertinência.

1.1 Objetivo

Visto que a lógica Clássica pode falhar na interpretação de problemas do mundo real, e os FIS têm sido amplamente utilizados para classificação de alimentos, juntamente com os algoritmos de AM, o objetivo desse trabalho é verificar e comparar a viabilidade do uso de modelos Fuzzy e métodos de AM aplicados a problemas de classificação que contém amostras não-factíveis. Essa viabilidade pode ser entendida como a análise do desempenho comparados entre os modelos Fuzzy e da lógica Clássica, desempenho esse medido pela quantidade de amostras classificadas. Além de utilizar técnicas de modelagem Fuzzy (*Top-Down*) e algoritmos de AM (*Bottom-Up*) com a finalidade de comparar acurácias e indicar a mais adequada para dado problema. Para isso, foram utilizados Árvore de Decisão J48, k-NN, MLP, SVM, Naïve Bayes (NB), *Random Forest* (RF) e FRBS.

1.2 Organização do Trabalho

O trabalho está organizado como segue:

- No Capítulo 2 estão descritos conceitos nos quais este trabalho foi baseado. São apresentados conceitos em relação a Modelagem Preditiva, Sistemas Baseados em Regras, Factibilidade, Aprendizado de Máquina, Fundamentos Fuzzy, Sistemas de Inferência Fuzzy e como eles podem ser utilizados para classificação.
- No Capítulo 3 é demonstrado o processo de inferência fuzzy para classificação, aplicado ao problema de classificação de carne suína quanto à qualidade.
- No Capítulo 4 são apresentados os trabalhos relacionados, primeiramente à classificação de produtos agrícolas com o uso de lógica fuzzy e, em seguida, abordagens para classificação de carnes suínas.
- No Capítulo 5 estão descritos o método de avaliação de qualidade de carne suína convencional e os padrões de avaliação utilizados, as bases de dados reais utilizadas, o software e os pacotes utilizados nos experimentos e os modelos obtidos.
- No Capítulo 6 são apresentados os experimentos e os resultados obtidos. Além das comparações entre os métodos e a discussão dos resultados.
- No Capítulo 7 são apresentadas as conclusões do trabalho por meio dos resultados, com discussões e trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Nesse capítulo serão apresentados os conceitos utilizados na realização do trabalho. A Seção 2.1 expõem conceitos de modelagem preditiva, conjuntos de dados, amostras e atributos, além de regressão e principalmente de classificação. Em seguida, na Seção 2.2 é apresentada a definição de factibilidade. A Seção 2.3 apresenta conceitos de Aprendizado de Máquina, entre tipos de aprendizado, algoritmos de classificação e balanceamento de dados. Na Seção 2.4 é apresentada a representação de Sistemas Baseados de Regras para classificação e a Seção 2.5 explica a Teoria de Conjuntos e Lógica Fuzzy, e como utilizar Sistemas de Inferência Fuzzy para classificação.

2.1 Modelagem Preditiva

Modelagem preditiva é uma das tarefas mais comuns da Mineração de Dados [1]. Como o próprio nome implica, é o processo de tomar dados históricos (datasets), identificar padrões nos dados que são vistos através de alguma metodologia (o modelo), e então utilizar o modelo para realizar previsões sobre o que irá acontecer no futuro.

De acordo com Faceli, “Um algoritmo de aprendizado de máquina preditivo é uma função que, dado um conjunto de exemplos rotulados, constrói um estimador” [32], em que o valor atribuído ao rótulo pertence a um domínio conhecido. Se o domínio desse rótulo for um conjunto infinito e ordenado de valores, denomina-se um problema de regressão, e se o domínio for um conjunto de valores nominais, tem-se um problema de classificação.

Dado um problema de classificação com duas classes, ou rótulos, o algoritmo tem o objetivo de encontrar uma fronteira de decisão que separe as amostras de cada classe. Caso as amostras sejam linearmente separáveis, como na Figura 1a, a fronteira de decisão pode ser definida por uma reta (podendo ser as retas A, B ou C), se não forem como na Figura 1b, será necessário combinar mais retas. Se o problema possuir mais de duas variáveis de entrada, será necessário utilizar hiperplanos de separação.

Em alguns casos, as amostras apresentam valores de atributos não-factíveis a uma dada classificação, por exemplo, valores fora do intervalo esperado. Nesse caso, a fronteira de decisão tem a sua delimitação comprometida. Os detalhes com relação a factibilidade serão abordados na Seção 2.2.

As fronteiras de decisão encontradas podem variar de acordo com o algoritmo de aprendizado de máquina. Além de outros fatores, como diferentes conjuntos de treinamento e variações na ordem de apresentação dos exemplos podem fazer com que o mesmo algoritmo encontre fronteiras diferentes a cada novo treinamento.

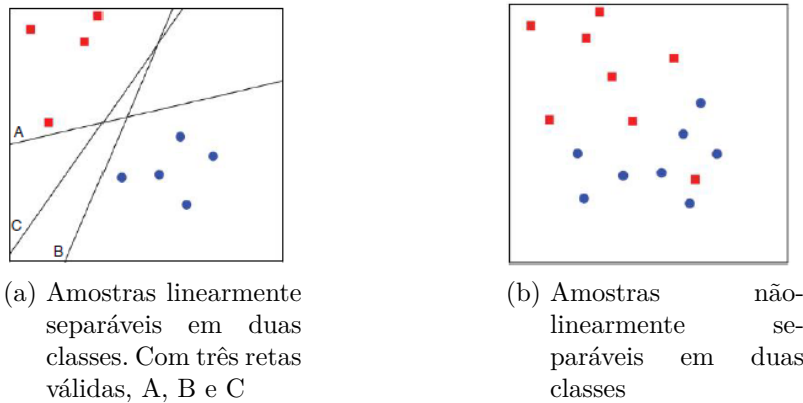


Figura 1 – Exemplo de amostras linearmente separáveis e não-linearmente separáveis [1]

Um dataset para classificação apresenta exemplos como os da Tabela 1, em que a última coluna é a classe ou atributo alvo, que assume valores discretos. Cada linha representa uma amostra e as demais colunas são os atributos de cada amostra de entrada.

Tabela 1 – Exemplo de dataset para classificação

pH final	L*	CRA	Classe
5,34	55,3000	3,7167	PFN
5,53	54,7333	3,4291	PFN
5,64	55,4333	2,4225	PFN
5,80	47,5667	2,0429	RFN
5,87	47,7667	2,2015	RFN
5,53	47,6667	2,7965	RFN

Métodos preditivos de AM podem ser organizados em baseados em distância como o k-NN, probabilístico como o classificador NB, os baseados em procura como as árvores de decisão e regressão, e os sistemas baseados em regras ou os métodos baseados em otimização, como as Redes Neurais Artificiais (RNAs) e as SVMs.

2.2 Factibilidade

A factibilidade pode se referir a possibilidade lógica, aquilo que pode ser alcançado, realizável. A possibilidade lógica se refere a uma proposição que pode ser um consequente lógico de outra, baseado em axiomas de um dado sistema lógico [33].

A possibilidade lógica de uma proposição depende do sistema lógico a ser considerado, não apenas pela violação de uma única regra. Alguns sistemas lógicos restringem inferências de proposições inconsistentes ou permitem contradições. Outros tem mais de um valor verdade em vez de valores binários. O critério para a possibilidade lógica é muitas vezes baseado em se a proposição é contraditória ou não.

Além disso, a possibilidade lógica deve ser diferenciada da possibilidade metafísica. Pois a primeira tem relação com o que é permitido em um sistema lógico e a segunda é relacionado ao que pode ser verdade em algum universo possível. Dessa forma, algumas proposições podem ser logicamente possíveis segundo um sistema lógico, mas serem metafisicamente impossível e ainda serem verdade.

Como por exemplo, um conjunto Fuzzy com as variáveis “criança”, “adolescente” e “adulto”, para determinar a fase da vida de uma pessoa de acordo com a idade dela. Uma amostra de entrada pode estar na interseção entre “criança” e “adolescente”, o que é possível logicamente e metafisicamente, pois pode ser uma pessoa na idade de transição. Porém, dependendo como foram limitados esses conjuntos, determinada entrada poderia estar na interseção entre “criança” e “adulto”, podendo ser logicamente possível dentro daquele sistema, porém, metafisicamente não faz sentido uma pessoa ser considerada criança e adulta ao mesmo tempo, de acordo com a sua idade.

As amostras não-factíveis podem simplesmente ser descartadas do dataset de análise, como fez Lira [34], que excluiu 178 amostras não-factíveis, pois não tinham as informações suficientes. Ou então, podem ser avaliadas e convertidas para factíveis, quando possível. Como em [35], em que existe um vetor de seleção binário sujeito a restrições, sendo possível aplicar uma operação extra para correção.

Nesse trabalho, são consideradas não-factíveis as amostras dos datasets reais que não atendem as restrições da lógica Clássica, ou seja, as amostras que a lógica Clássica não foi capaz de rotular.

2.3 Aprendizado de Máquina

Na maioria dos problemas de aprendizado, a tarefa é aprender para classificar entradas de acordo com um conjunto finito (as vezes infinito) de classificações. Geralmente, é fornecido um conjunto de dados classificados à mão para o treinamento do sistema de aprendizado [6]. Então, o sistema tenta aprender a partir dessas amostras de treinamento, a fim de generalizar as informações e ser capaz de classificar novas amostras a partir do conhecimento obtido.

Diz-se que um agente está aprendendo se ele melhora seu desempenho em tarefas futuras após observar o mundo [36]. Um agente precisa aprender pois não é possível o projetista antecipar todas as situações possíveis e passar esse conhecimento para ele, como um robô projetado para navegar em labirintos precisa aprender a configuração deles a cada novo encontro. Como também os projetistas não podem antecipar todas as mudanças ao longo do tempo ou mesmo quando eles não sabem como programar uma solução.

Existem três principais tipos de aprendizado, o supervisionado, o não-supervisionado e o semi-supervisionado. No aprendizado supervisionado o agente observa alguns exem-

plos de pares de entrada-saída e aprende uma função que mapeia o valor de entrada com o valor de saída.

No aprendizado não-supervisionado o agente aprende padrões de entrada sem o fornecimento de um *feedback* explícito, ele não tem conhecimento da relação dos valores de entrada e saída. Em contrapartida, existe o aprendizado semi-supervisionado, em que dado alguns exemplos rotulados é necessário que o agente resolva uma quantidade de amostras não rotuladas. Na prática, essa distinção nem sempre é muito bem definida [36]. Neste trabalho, foram utilizados oito métodos distintos de aprendizado de máquina para classificação. A seguir, são descritos seus principais conceitos.

- Sistemas Fuzzy Baseados em Regras (FRBS): método bastante conhecidos em *soft-computing*, baseados em conceitos fuzzy [37]. Tornaram-se métodos poderosos pois lidam com diversos problemas, como incerteza, imprecisão e não-linearidade. Os FRBSs são uma extensão dos sistemas baseados em regras clássico. Nos experimentos foram utilizados o Fuzzy-Chi, baseado na distribuição Chi e o Fuzzy-W, que é um algoritmo fuzzy com pesos e aplica a estratégia de Ishibuchi [38].
- Árvore de Decisão J48 (J48): os algoritmos de árvore de decisão estão entre os mais comuns e poderosos métodos na área de aprendizado de máquina [39], são computacionalmente baratos, lidam com atributos irrelevantes [40] e apresentam resultados de fácil entendimento. O algoritmo J48 é uma implementação otimizada em Java do algoritmo C4.5 [41]. Esse algoritmo explora o conceito de ganho de informação de cada atributo, separando os dados em planos ortogonais no espaço de saída. A ideia principal é decompor um problema complexo em sub-problemas mais simples, aplicando o mesmo critério recursivamente nos sub-problemas até o critério de parada ser satisfeito.
- Algoritmo de Agrupamento *k-Nearest Neighbor* (k-NN): entre os métodos baseados em instâncias, o k-NN é considerado o mais simples [42], apresenta alta acurácia e não é sensível a *outliers* [40]. Este algoritmo assume que as amostras correspondem a ponto em um espaço de n -dimensões (sendo n a quantidade de atributos). Dada uma nova amostra, seus k vizinhos mais próximos são definidos em termos da distância, normalmente a Euclidiana, entre todos os atributos da amostra. Ao contrário de métodos como as árvores de decisão, que selecionam um subconjunto de características [43]. E então, o valor predito para a amostra de entrada é a média entre os vizinhos.
- *Multilayer Perceptron* (MLP): as redes multicamadas *feed-forward* são uma classe importante de redes neurais [44]. Tipicamente, a rede consiste de: um conjunto de neurônios que constituem a camada de entrada; uma ou mais camadas ocultas intermediárias; e uma camada de saída. O sinal de entrada se propaga através da rede em

uma direção, camada por camada. Essas redes neurais são conhecidas como MLPs. Elas têm sido aplicadas com sucesso para resolver problemas diversos utilizando o algoritmo de *backpropagation*, devido a sua natureza de aprendizado neural.

- *Naïve Bayes* (NB): o NB é um classificador probabilístico simples, amplamente estudado desde os anos 50 [42]. Ele é capaz de lidar com múltiplas classes [40], tem seu desempenho comparável com árvores de decisão e redes neurais em alguns domínios e é frequentemente utilizado em *benchmarks* como comparação para novos classificadores [45].
- *Random Forest* (RF): a RF é um conjunto de várias árvores de decisão (*ensemble*), proposto por Breiman [46]. Essas árvores são construídas de forma independente, sem poda e realizam a classificação através de votação. A quantidade de árvores e atributos são hiper-parâmetros definidos pelo usuário. Além disso, esse método é baseado em *bagging* [47], em que cada árvore utiliza um subconjunto de treinamento com amostras aleatórias, o que aumenta sua capacidade de generalização.
- *Support Vector Machine* (SVM): SVM é um método de classificação do estado da arte introduzido por [48], que pertence a categoria de métodos baseados em núcleos (*kernels*). As SVM's são amplamente utilizadas no campo de aprendizado de máquina devido a sua alta acurácia, habilidade de lidar com dados de alta dimensão e flexibilidade de modelagem diversos tipos de dados [49].

No entanto, diversos problemas do mundo real, lidam com a distribuição não uniforme das classes dos dados, podendo existir mais amostras de uma classe do que a outra, em que, a classe com mais amostras é denominada majoritária e a segunda, minoritária [50]. Por exemplo, as amostras positivas em um dataset sobre câncer serão bem menores do que as negativas ou o número de amostras de fraudes, que serão menores do que a quantidade de amostras de transações não fraudulentas. Esse desbalanceamento causa a redução da acurácia da predição das amostras da classe minoritária mas torna a acurácia geral alta. Além de que, ignorar a taxa de classificação errada pode causar problemas graves, como em casos de diagnóstico médico, por exemplo.

A re-amostragem pode ser a solução para esses casos [51]. Pois os métodos de re-amostragem são utilizados para alterar a proporção da distribuição das classes para balancear o dataset. Eles podem ser considerados de duas formas: sobre-amostragem ou sub-amostragem. A primeira é o aumento do número de amostras da classe minoritária e a segunda é a redução do número de amostras da classe majoritária.

Dessa forma, como os datasets com amostras reais utilizados nesse trabalho não atendem às restrições necessárias, os algoritmos de aprendizado de máquina utilizaram

dataset sintéticos criados a partir dos padrões para a comparação dos resultados. Os datasets serão discutidos com mais detalhes na Seção 5.2.

2.4 Sistemas Baseados em Regras

Uma forma de representar o conhecimento é através de regras que expressam o que pode acontecer ou o que acontece quando determinadas condições são atendidas [6]. Essas regras normalmente são expressas na forma de sentenças *SE A ENTÃO B*. Em que A é o antecedente da regra e B o conseqüente, que geralmente é uma ação ou conclusão. Em outras palavras, o objetivo da regra é informar ao sistema qual decisão tomar baseado nas circunstâncias ou a qual conclusão chegar a partir de um conjunto de amostras de entrada.

De forma geral, uma regra pode ter mais de um antecedente, usualmente combinados por operadores lógicos E ou OU. Como também podem ter mais de um conseqüente, o que sugere que existem múltiplas ações a serem tomadas.

A parte antecedente ou o condicional A é um conjunto de condições, em que cada condição é definida por uma relação entre um atributo e os valores do domínio [32]. Essa relação pode ser =, <, >, ≥, ≤ ou ∈, como por exemplo: *SE x > 5*; *SE nome = Maria*; *SE clima = frio*. Dessa forma, os atributos são *x*, *nome* e *clima*, os operadores são > e = e os valores são 5, *Maria* e *frio*.

Sistemas Baseados em Regras (*Rule-Based System* - RBS) utilizam regras para fornecerem sugestões, diagnósticos, determinar uma ação em uma situação particular ou resolver problemas específicos. Um RBS consiste de componentes como uma base de regras, também chamada de base de conhecimento, uma base de fatos e um interpretador ou motor de inferência.

A base de conhecimento é um conjunto de regras que representam o conhecimento do sistema. A base de fatos representa as entradas do sistema, que são utilizadas para chegar a conclusões ou tomar ações. O motor de inferência é a parte do sistema que controla os processos de tomada de decisão a partir da combinação das regras e os fatos.

RBS tem o objetivo de permitir o uso das mesmas regras que os especialistas utilizam para tirar conclusões do conjunto de dados que são apresentados para o sistema, ou seja, o sistema utiliza o julgamento humano para tomar decisões, o que agiliza o processo realizado por eles. A grande vantagem da representação para generalização de exemplo com regras de decisão é que cada regra é independente das outras, e pode ser interpretada de forma isolada.

Existem duas estratégias básicas para encontrar uma regra. A primeira é denominada *Top-Down* e é orientada pelo modelo, ela parte da regra mais geral, aplica operadores

de especificação e acrescenta condições à parte condicional da regra. A segunda é chamada *Bottom-Up* e é orientada pelos dados, ao contrário da primeira, ela inicia pela regra mais específica, aplica operadores de generalização e remove restrições [36, 32].

2.5 Fundamentos Fuzzy

Ao contrário da lógica Aristotélica em que uma proposição apenas pode tomar valores verdadeiro ou falso como verdade, sempre um ou outro, nunca ambos nem nenhum, a teoria de conjuntos Fuzzy e a lógica Fuzzy introduzem o conceito de grau de pertinência. Ele é um valor real entre 0 e 1, que indica o quão verdadeira uma proposição pode ser ou quanto um elemento pertence ao conjunto.

A teoria de conjuntos Fuzzy é uma generalização da teoria de conjuntos convencional e foi introduzida por Zadeh em 1965 [7]. Ela fornece uma ferramenta matemática para lidar com variáveis linguísticas associadas com linguagem natural.

As variáveis linguísticas representam os atributos linguisticamente, elas têm um significado semântico, podem assumir um valor em um intervalo de valores fuzzy [6, 52] e são um componente muito importante e bastante utilizada na teoria de conjuntos Fuzzy e na lógica Fuzzy. Por exemplo, a variável linguística para representar o conceito de “altura”, por exemplo, que pode ser representada por intervalos, denominados termos linguísticos, como “alto”, “médio” e “baixo”. Ela pode ser definida no universo de 0,60 m à 2,40 m, e seus termos linguísticos definem os subconjuntos desse universo.

2.5.1 Conjuntos Fuzzy

Na teoria de conjuntos *crisp*, um conjunto pode ser definido pelos valores que estão contidos dentro dele, dessa forma, o valor pertence ou não ao conjunto. Como por exemplo, o conjunto de números naturais $N = \{0, 1, 2, 3, 4, \dots\}$ é um conjunto *crisp*, em que esse valores definitivamente pertencem a ele. Valores como 0,5 e π não são membros desse conjunto.

Em contrapartida, ao considerar o conjunto fuzzy de pessoas altas, todos pertencerão a ele, cada um com seu grau de pertinência ao conjunto, podendo ser alto e não alto, até certo ponto. Como por exemplo, João tem 2,15 m de altura, Maria tem 1,25 m e Helena tem 1,60 m. João com certeza está incluso no conjunto de pessoas altas com grau de pertinência igual a 1, Maria tem um grau de pertinência igual a zero, e o grau de pertinência de Helena estaria perto de 0,5.

Portanto, a definição de um conjunto fuzzy é uma maneira mais natural de expressar como as pessoas realmente falam sobre as coisas. É mais comum alguém dizer que uma pessoa é “muito alta” ou “não muito alta” do que apenas “alta” e “não alta”. Isso

acontece pois cada pessoa tem sua própria concepção quanto ao conceito de “alto” e não é precisa.

A base da lógica Clássica é que uma proposição precisa ser necessariamente, ou verdadeira ou falsa, ela não pode ser verdadeira e falsa ao mesmo tempo, e também não é possível não ser verdadeira nem falsa. Por outro lado, na lógica Fuzzy uma proposição pode ser verdadeira e falsa ao mesmo tempo, como pode não ser verdadeira nem falsa.

2.5.1.1 Funções de Pertinência

A escolha e a modelagem das funções de pertinência dos conjuntos fuzzy, tem um papel importante na teoria de conjuntos fuzzy e suas aplicações. As funções de pertinência podem ser Gaussiana, Sino, Sigmóide, Trapezoidal ou Triangular e a escolha entre esses tipos varia de acordo com o contexto ou especialista.

As funções Gaussiana, Trapezoidal e Triangular, por exemplo, são definidas pelas Equações 2.2, 2.3 e 2.1, respectivamente. Em que, p_1 , p_2 , p_3 , e p_4 são os pontos que indicam os vértices para construir a função. O intervalo do grau de pertinência sempre varia entre 0 e 1, e o intervalo dos valores de entrada varia de acordo com a modelagem.

$$f(x) = \begin{cases} 0 & \text{se } x \leq p_1 \\ 0 & \text{se } x \geq p_2 \\ 1 = m(p_2 - p_1) & \text{se } p_1 > x > p_2 \\ 1 = m(p_3 - p_2) & \text{se } p_2 > x > p_3 \end{cases} \quad (2.1)$$

$$g(x, \sigma, c) = e^{-\frac{(x-c)^2}{2\sigma^2}}, \text{ em que } \sigma = \text{variância e } c = \text{média} \quad (2.2)$$

$$h(x) = \begin{cases} 0 & \text{se } x \leq p_1 \\ 0 & \text{se } x \geq p_4 \\ 1 & \text{se } p_2 \geq x \geq p_3 \\ 1 = m(p_2 - p_1) & \text{se } p_1 > x > p_2 \\ 1 = m(p_4 - p_3) & \text{se } p_3 > x > p_4 \end{cases} \quad (2.3)$$

A Figura 2 ilustra como é calculado o grau de pertinência de uma amostra de entrada, com a utilização das funções Gaussiana, Trapezoidal e Triangular. Na Figura 2a pode ser visualizada a função de pertinência do tipo Gaussiana construída com $\sigma = 0,2$ e $c = 0,5$, a função Triangular na Figura 2c com $p_1 = 0,3$, $p_2 = 0,5$ e $p_3 = 0,7$ e a função Trapezoidal na Figura 2b com $p_1 = 0,3$, $p_2 = 0,4$, $p_3 = 0,6$ e $p_4 = 0,7$, com os dois eixos variando de 0 a 1.

Para o exemplo da Figura 2, o mesmo valor de entrada 0,5 é utilizado como amostra para o atributo x para as três funções. Esse valor é mapeado na função e seu grau de pertinência é dado por $MF.grau(x)$. Dessa forma, o grau de pertinência da amostra foi igual a 1 (valor máximo), pois alcançou o ponto mais alto em todas as funções. Porém, é possível perceber que uma pequena variação do valor de entrada, mudaria o grau de pertinência resultante das funções do tipo Gaussiana e Triangular, enquanto a função do tipo Trapezoidal abrange mais valores correspondentes ao grau de pertinência máximo.

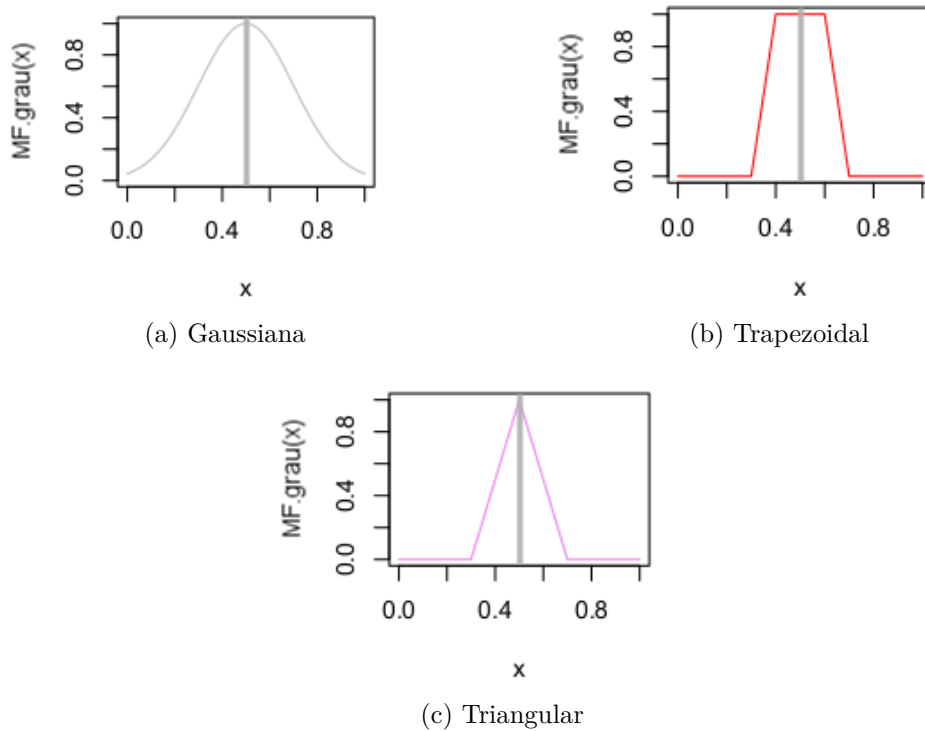


Figura 2 – Entrada $x = 0,5$ para as diferentes formas de funções de pertinência. (a) Gaussiana com $\sigma = 0,2$ e $c = 0,5$. (b) Trapezoidal com $p_1 = 0,3$, $p_2 = 0,4$, $p_3 = 0,6$ e $p_4 = 0,7$. (c) Triangular com $p_1 = 0,3$, $p_2 = 0,5$ e $p_3 = 0,7$.

Um conjunto fuzzy A é definido de acordo com suas funções de pertinência, M_A . Por exemplo, para o conjunto da altura das pessoas como ilustrado na Figura 3, o eixo x varia de 0,6 m a 2,4 m para valores de entrada para a altura e o eixo y varia de 0 a 1 para valores do grau de pertinência.

A função de pertinência para o termo linguístico “baixo” tem intervalo de 0 a 1,5 com maior grau de pertinência entre 0 e 1,4. Para o termo “médio” a base da função varia de 1,45 a 1,8, com maior grau de pertinência entre 1,55 e 1,75. Para o termo “alto” o intervalo varia entre 1,7 e 2,4, com valores máximos de pertinência de 1,9 a 2,4. Ou seja, para uma pessoa ser considerada definitivamente alta, de acordo com essas funções de pertinência, ela precisa medir entre 1,9 m e 2,4 m, qualquer outro valor entre 1,7 m e 1,8 m ela pode ser considerada de estatura nem alta nem média, ou média e alta ao mesmo tempo, por exemplo.

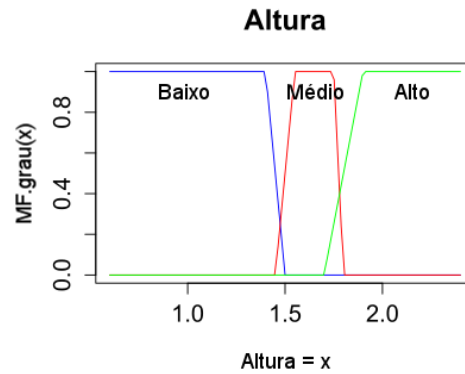


Figura 3 – Função de Pertinência para o Conjunto de Altura de Pessoas

2.5.1.2 Operadores de Conjuntos

Diversos operadores da teoria de conjuntos tradicional também podem ser aplicados nos conjuntos A e B , como o complemento de A que contém os elementos que não estão contidos em A ; a interseção entre A e B que contém apenas os elementos que estão contidos em A e B ao mesmo tempo; e a união entre A e B que contém todos os elementos contidos em A e todos os elementos contidos em B .

Pode-se relacioná-los com os operadores lógicos \neg (negação), \wedge (e) e \vee (ou). Dessa forma, os operadores seguem as leis de DeMorgan e são comutativos, associativos e distributivos. Para os conjuntos fuzzy, operadores semelhantes podem ser definidos.

Para exemplificar, ao realizar a interseção dos conjuntos de pessoas de estatura média (M) e alta (A), a pessoa que estiver no conjunto $M \cap A$ ela é considerada tanto média quanto alta. Logo, a interseção de dois conjuntos fuzzy é o conjunto dos elementos que pertencem a ambos [53]. Da mesma maneira, uma pessoa que pertence ao conjunto $M \cup A$ é ou média ou alta.

2.5.2 Sistemas de Inferência Fuzzy

A lógica fuzzy é uma forma de lógica que se aplica a variáveis fuzzy [53]. Cada variável fuzzy pode ter valor de 0 (definitivamente não verdadeiro) a 1 (definitivamente verdadeiro), como também pode assumir valores reais entre 0 e 1, em que 0,5 pode significar “pouco verdadeiro” ou “tanto verdadeiro quanto falso”. Além de que, os operadores de conjunto fuzzy podem ser aplicados nos cálculos com variáveis fuzzy.

Um tipo de inferência bastante utilizado em sistemas fuzzy é a Mamdani [6], que foi criada pelo professor Ebrahim Mamdani na década de 1970 para controlar um motor a vapor e uma cadeira. Ela permite que o sistema receba um conjunto de valores *crisp* como entrada e aplica um conjunto de regras fuzzy nesses valores, a fim de obter um único valor de saída *crisp*.

Dessa forma, um FIS pode ser construído a partir de variáveis linguísticas apro-

priadas ao problema e pela definição de suas funções de pertinência. Além da geração das regras baseadas no conhecimento do especialista com o uso das variáveis linguísticas. Em seguida, as regras são aplicadas no conjunto de entrada, com o método de inferência, a fim de obter um valor de saída [6], como ilustrado na Figura 4.

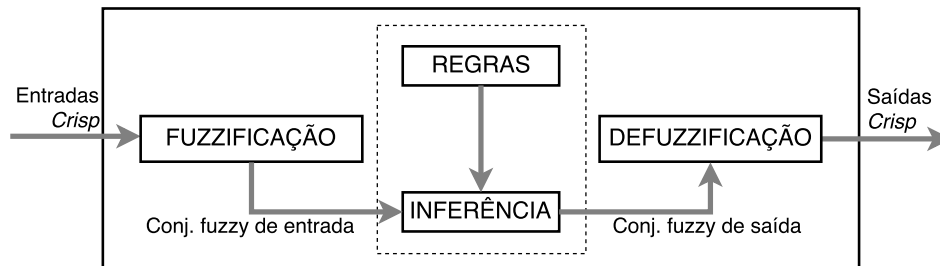


Figura 4 – Sistema de Inferência Fuzzy. Adaptado de [2]

As variáveis linguísticas fuzzy são representações linguísticas dos atributos. Elas tem um significado semântico e podem ser representadas numericamente por suas funções de pertinência [52]. Por exemplo, a variável linguística $pH(x)$, para a representação do atributo pH. Dessa forma, $pH(x) = \{baixo, medio, alto\}$ pode ser o conjunto de decomposições para esta variável linguística, e cada membro desse conjunto é denominado termo linguístico.

Funções de pertinência são utilizadas para quantificar um termo linguístico. Elas estão presentes nos passos de Fuzzificação e Defuzzificação, para mapear os valores de entrada *crisp* para termos linguísticos fuzzy e valores fuzzy para valores de saída *crisp*.

A base de regras fuzzy é uma coleção de regras SE-ENTÃO (IF-THEN) que controlam a variável de saída. Elas podem ser extraídas dos dados numéricos ou fornecidas por especialistas. As regras Zadeh-Mamdani podem ser expressas como:

$$IF\ x\ is\ A,\ THEN\ y\ is\ B$$

em que $(x\ is\ A)$ e $(y\ is\ B)$ são proposições fuzzy, x e y são variáveis definidas sobre os conjuntos universo U e V , respectivamente, e a parte “ $x\ is\ A$ ” é o antecedente e “ $y\ is\ B$ ” é o conseqüente. A e B são conjuntos fuzzy definidos por suas funções de pertinência [52], como *baixo* e *PSE*. Por exemplo, nós temos *IF pH is baixo, THEN class is PSE*.

Os operadores fuzzy são aplicados para avaliar, combinar os resultados individuais das regras e reformular a função de pertinência de saída para cada uma. A Tabela 2 apresenta os operadores mais básicos de conjuntos fuzzy, a interseção (t-norm) e a união (s-norm) [54].

Após a aplicação dos operadores e do método de implicação, é realizada a defuzzificação dos conjuntos de saída fuzzy. A defuzzificação é o processo de calcular uma única saída numérica para uma variável de saída fuzzy baseado nas funções de pertinência resultantes da implicação. Dois métodos amplamente utilizados são o centro de gravidade

Tabela 2 – Operadores Fuzzy T-norm e S-norm

T-norm/E/Interseção	S-norm/OU/União
$T(a, b) = \min\{a, b\}$	$S(a, b) = \max\{a, b\}$
$T(a, b) = a * b$	$S(a, b) = a + b - a * b$
$T(a, b) = \max\{0, a + b - 1\}$	$S(a, b) = \min\{1, a + b\}$

que encontra o centro geométrico das funções de pertinência de saída e o cálculo da média do máximo, que encontra o maior grau de pertinência das saídas fuzzy, se houver mais de um valor com o grau máximo, é calculada a média entre eles.

3 PROCESSO DE INFERÊNCIA FUZZY

Nesse capítulo será apresentado um exemplo de acordo com o diagrama do FIS da Figura 4 para o problema de classificação de carne suína, seguindo a inferência Mamdani e o padrão de qualidade de Warner [5]. O processo consiste nas etapas de 1) Fuzzificação do conjunto de entradas *crisp*; 2) Inferência do antecedente para o conseqüente, de acordo com o conjunto de regras da base de conhecimento; 3) Defuzzificação dos conjuntos de saída fuzzy para saídas *crisp*.

De forma geral, o FIS funciona da seguinte maneira: dado um conjunto de entradas *crisp* para o sistema, elas são fuzzificadas, ou seja, são transformadas em subconjuntos fuzzy através do mapeamento em funções de pertinência para o cálculo de seus graus de pertinência para cada parte do antecedente da regra; após a fuzzificação das entradas, os operadores fuzzy são aplicados para determinar o ponto em que as funções de pertinência de saída serão “redesenhadas”, pois seus limites superiores podem sofrer alterações.

O método de inferência (ou implicação), utilizará o resultado do operador fuzzy para reformular as funções de pertinência de saída, e o resultado pode ser diferente de acordo com o método de implicação utilizado. Como o sistema precisa retornar um valor *crisp*, então é realizada a defuzzificação e agregação dos resultados (pode ser ao contrário também, como será discutido ao longo do capítulo).

A defuzzificação é a transição do conjunto fuzzy para um valor *crisp* e a agregação, como o nome diz, é a combinação dos resultados de cada regra, pois todos tem sua importância. E assim temos o resultado final do sistema.

Para o problema de classificação de amostras de carne suína quanto a sua qualidade, um dos padrões de avaliação foi proposto por Warner [5]. Esse padrão avalia medidas de pH (indica acidez ou basicidade de uma solução), L^* (medida de luminância do padrão CIELab) e capacidade de retenção de água (CRA) e classifica as amostras entre quatro classes. Os nomes das classes é um sistema de três letras que consiste de uma tripla relacionada a cor, firmeza e exsudação, são elas: PSE - pálida, macia e exsudativa (*pale, soft, exudative*); RSE - vermelha, macia e exsudativa (*red, soft, exudative*); RFN - vermelha, firme e não-exsudativa (*red, firm, non-exudative*); DFD - escura, firme e seca (*dark, firm, dry*). O padrão deixa de fora alguns valores como 42 e 49 para L^* , por exemplo. Na Tabela 3 podem ser visualizados os intervalos dos atributos para cada classe.

Para esse exemplo foi escolhida uma amostra de classe RFN, com $pH = 5,56$, $L^* = 47,9$, e $CRA = 2,51$. Nas próximas seções serão demonstrados passo-a-passo como são realizadas as etapas desse processo. A Seção 3.1 demonstra como uma entrada *crisp* é fuzzificada, a Seção 3.2 demonstra a aplicação do operador fuzzy nos graus de pertinência

Tabela 3 – Padrão Warner [5]

Classe	pH	L*	CRA
PSE	<6	>50	>5
RSE	<6	42-50	>5
RFN	<6	42-50	<5
DFD	≥ 6	<42	<5

do antecedente, e então a implicação do antecedente no consequente na Seção 3.3 e por último, na Seção 3.4 a etapa de defuzzificação e agregação dos conjuntos de saída fuzzy para as saídas *crisp* do sistema.

3.1 Passo 1: Fuzzificar as Entradas do FIS

O primeiro passo é determinar os graus de pertinência de cada entrada *crisp* para cada conjunto fuzzy apropriado via função de pertinência. Esse exemplo foi construído com quatro regras, e cada uma delas depende da resolução das entradas para uma quantidade diferente de conjuntos linguísticos fuzzy: *pH é baixo*, *pH é alto*, *L* é baixo*, *L* é médio*, *L* é alto*, *CRA é baixo* e *CRA é alto*.

Antes da avaliação das regras, é necessário que as entradas *crisp* sejam fuzzificadas de acordo com as variáveis e termos linguísticos. Dada amostra de entrada, seus valores *crisp* são mapeados simultaneamente nas funções de pertinência dos respectivos atributos, de forma simultânea entre todas as regras. Por exemplo, a Figura 5 ilustra o valor de entrada para o pH igual a 5,56. Como a variável linguística do pH apresenta dois termos linguísticos, temos que essa amostra pertence 0,59 ao termo “baixo” e 0,41 ao termo “alto”. Então esses são os resultados da fuzzificação para o valor de pH da amostra escolhida.

3.2 Passo 2: Aplicar o Operador Fuzzy

Após a fuzzificação das entradas, sabe-se o grau em que é satisfeita cada parte do antecedente para cada regra. Se o antecedente de uma dada regra tem mais de uma condição a ser satisfeita, o operador fuzzy é aplicado para obter um valor que representa o resultado do antecedente para aquela regra. O valor obtido será usado como entrada para o método de implicação do antecedente no consequente. A entrada para o operador fuzzy são os graus de pertinência da parte antecedente da regra.

No exemplo, foi utilizada a função MIN para o operador t-norm (AND). A Figura 6 mostra a aplicação do operador em uma das regras da base de conhecimento construídas a partir do padrão Warner. Como pode ser observado, os graus de pertinência de $L^*=47,9$ e $CRA=2,51$ são iguais a 1 para L^* médio e CRA baixo, respectivamente, e 0,59 para

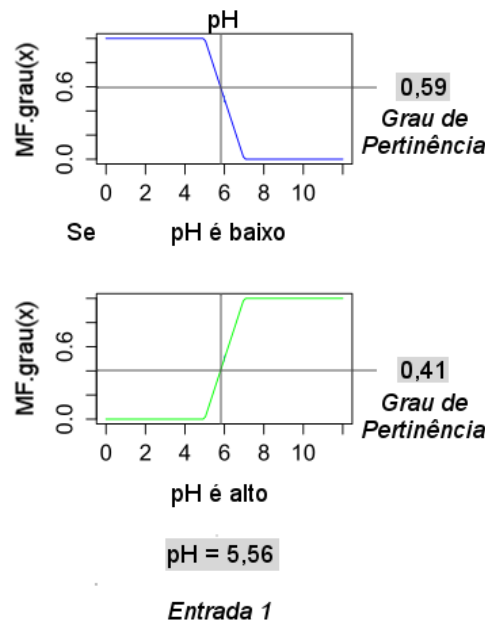


Figura 5 – Fuzzificação para o pH

pH baixo, como visto anteriormente. O operador AND faz a interseção e a função MIN, $\min(x_1; x_2)$, escolhe o menor valor e a operação fuzzy para regra completo. Outras opções como PRODUCT ($x_1 * x_2$) ou BOUNDED $\max(0, x_1 + x_2 - 1)$ também poderiam ser escolhidas.

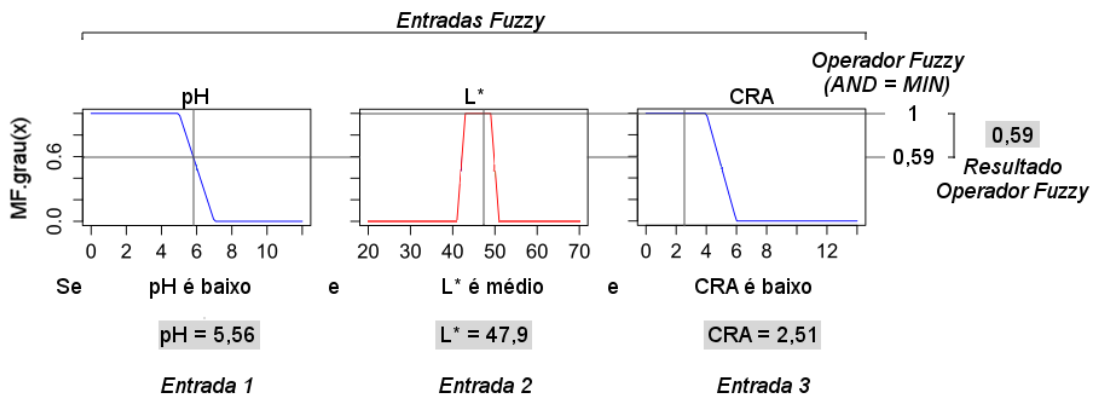


Figura 6 – Aplicação do Operador Fuzzy AND na Regra

Se no lugar do operador AND fosse escolhido o operador OR, o operador MAX $\max(x_1; x_2)$ poderia ser aplicado e o resultado seria 1 em vez de 0,59. No exemplo, o operador AND foi estabelecido devido às condições das regras do padrão de classificação de carne suína escolhido e a operação MIN por ser a mais utilizada [52].

3.3 Passo 3: Aplicar Método de Implicação

O consequente é um conjunto fuzzy representado por uma função de pertinência, que representa de forma adequada as características linguísticas que são atribuídas a ele. O consequente é reformulado utilizando uma função associada ao antecedente. A entrada para o processo de implicação é o resultado dado pelo antecedente (o grau de pertinência calculado de acordo com o operador fuzzy) e a saída do processo é um conjunto fuzzy.

A Figura 7 demonstra como é aplicado o processo de implicação na regra com o método MIN que trunca o conjunto de saída fuzzy de acordo com o resultado do antecedente.

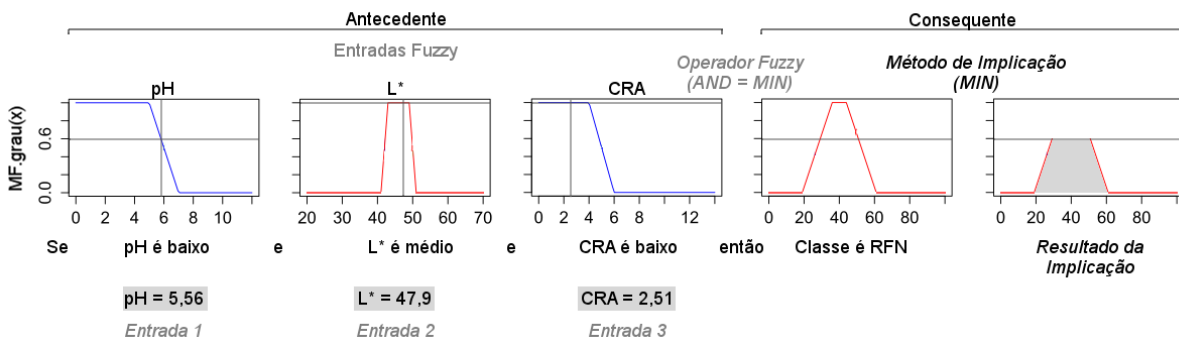


Figura 7 – Aplicação do Método de Implicação (MIN) na Função de Pertinência de Saída

3.4 Passo 4: Defuzzificação e Agregação

Como o sistema precisa dar uma saída *crisp* e, por enquanto, tem-se apenas conjuntos fuzzy, o processo de Defuzzificação precisa desenvolver a tarefa de agregar as informações fornecidas por cada um dos conjuntos fuzzy e transformá-los em um único valor *crisp*. Para isso, existem duas formas de realizar a agregação [55]: 1) Primeiro a agregação depois a fuzzificação; 2) Primeiro a fuzzificação depois a agregação.

De acordo com a primeira forma, o conjunto de regras individuais são agregados para obter o conjunto fuzzy final. Através de um operador de agregação e do método de defuzzificação, o conjunto fuzzy é transformado em um valor *crisp*. Dessa forma, é realizada a agregação dos conjuntos de saída resultantes do passo anterior em um único conjunto fuzzy e então, nesse conjunto é aplicada a defuzzificação para obter a saída *crisp* do sistema.

Da segunda maneira, a contribuição de cada conjunto fuzzy é individualmente considerada pela aplicação da defuzzificação em cada conjunto e o valor *crisp* de saída é obtido pelo operador de agregação aplicado nos resultados da defuzzificação. Ao contrário do método anterior, a defuzzificação é executada para cada regra, tendo como entrada

os conjuntos fuzzy de saída e então a função de agregação utiliza todos os resultados da defuzzificação para calcular o valor *crisp* de saída.

Métodos de defuzzificação como centro de gravidade, média do valor máximo e de agregação como média ou a média ponderada podem ser utilizados. No exemplo foi utilizada a segunda abordagem (defuzzificação e depois agregação) com o cálculo da média ponderada. O valor *crisp* de saída obtido foi 36, que se mapeado de volta na função de pertinência de saída, corresponde a classe RFN, como o rótulo da amostra do exemplo.

Na Figura 8 pode ser observado o diagrama de inferência fuzzy completo para o problema exemplo. Em que, dado os valores de entrada *crisp* $pH=5,56$, $L^*=47,9$ e $CRA=2,51$, eles são mapeados, simultaneamente, nas respectivas funções de pertinência, fuzzificados e os graus de pertinência de cada um são determinados. A seguir é aplicado o operador fuzzy nos antecedentes das regras e depois a implicação do antecedente no consequente. A partir do método da implicação, as funções de pertinência de saída são reformuladas e então aplicadas as funções de defuzzificação nos conjuntos de saídas fuzzy. Porém, como a saída do sistema precisa ser um único valor *crisp*, é realizada a agregação dos resultados da defuzzificação de cada regra e obtém-se a saída do sistema.

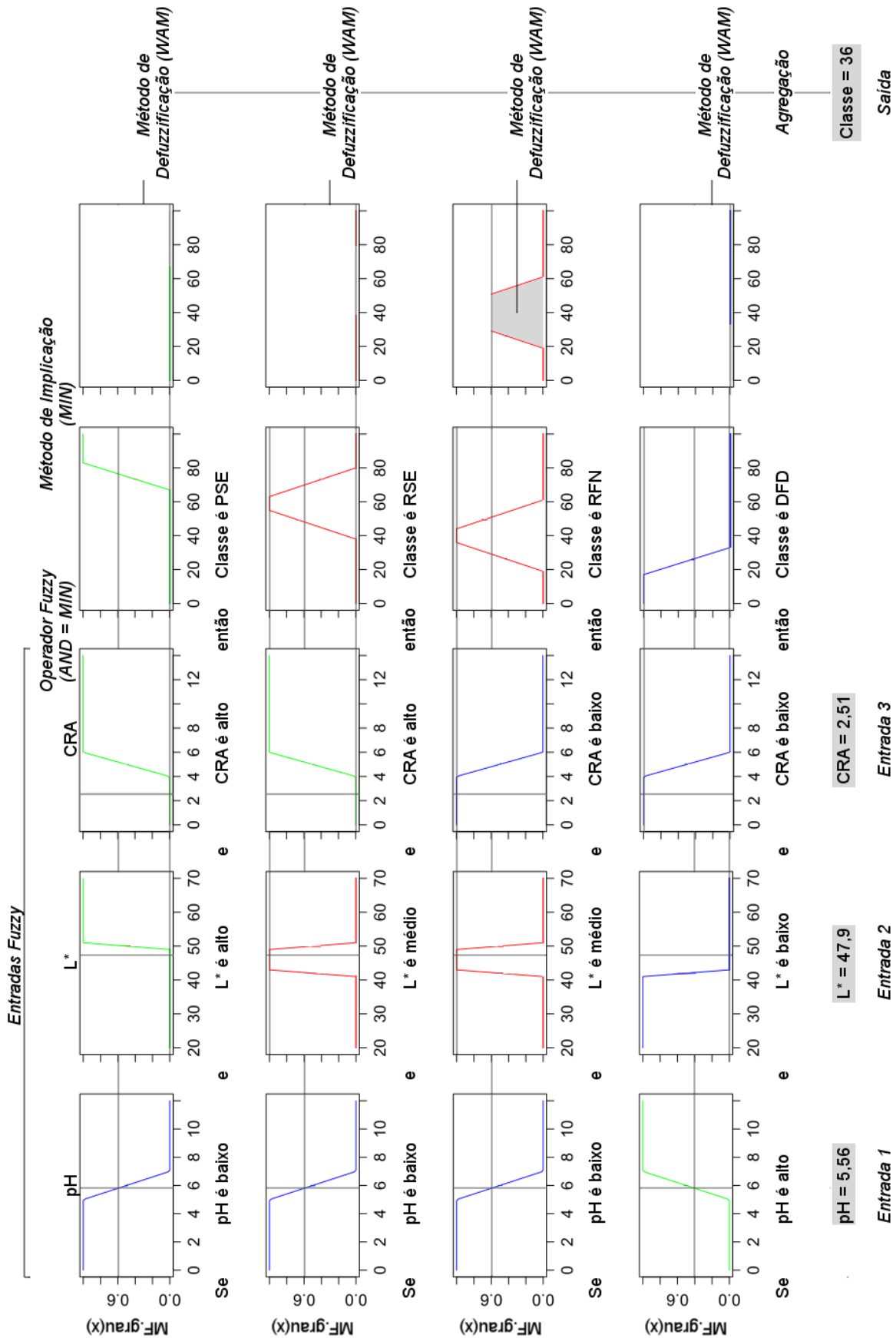


Figura 8 – Diagrama de Inferência Fuzzy, demonstrando a quantidade de regras criadas considerando o número de classes

4 TRABALHOS RELACIONADOS

O FIS é amplamente utilizado para classificação de produtos agrícolas, como maçãs, mangas [14, 15, 17], tâmaras [16] e tomates [3], por exemplo. Além do uso para classificação de carne suína [56, 57, 58, 59, 60, 61, 62], pois representa uma boa abordagem quando a experiência humana precisa ser incorporada no processo de tomada de decisão [15].

Kavdir *et al.* [14] classificaram 181 amostras de maçãs quanto a cor, defeitos, forma, peso e tamanho, em três classes (bom, médio e ruim). Ao comparar com a classificação feita anteriormente por humanos, obteve 89% de acurácia. Da mesma forma, Oathman *et al.* [17] obtiveram 87% de acurácia ao classificar mangas quanto a maturação, baseado apenas na cor. Foram 150 amostras classificadas previamente por humanos em três classes (não-maduro, maduro, passado), divididas em 50 amostras cada classe. Após utilizar SVM para reconhecer a frutas, entre maçãs, bananas, cenouras, mangas e laranjas, Mustafa *et al.* [15] utilizaram o FIS para classificação de maçãs, bananas e mangas quanto ao tamanho, entre *pequeno*, *médio*, *grande* e *extra grande*, e atingiu 98,85%, 98,75% e 89,74% de acurácia, respectivamente para cada fruta.

Para classificar 500 amostras de tâmaras quanto a sua qualidade, Avali *et al.* [16] utilizaram FIS com os atributos comprimento e frescor e cinco classes. A quantidade de termos linguísticos foram escolhidos de acordo com o especialista, as regras foram geradas a partir da combinação de todas as variáveis linguísticas e variou as formas das funções de pertinência para encontrar a melhor acurácia, que foi 91%.

Goel *et al.* [3] utilizaram FIS devido a sua necessidade de estimar a maturação de tomates não colhidos, de acordo com o grau de maturação do Departamento de Agricultura dos Estados Unidos (*United States Department of Agriculture* - USDA) ilustrado na Figura 9, sem que o método interferisse no crescimento natural das amostras. Para isso, foram avaliadas 116 imagens reais de acordo com a cor e classificadas em seis classes. As variáveis linguísticas foram escolhidas através de algoritmos de aprendizado e as regras foram geradas por árvores de decisão. A acurácia obtida foi de 94,29% com *holdout* estratificado, sendo 70% das amostras para treinamento e 30% para teste. O método fuzzy foi escolhido devido a sua capacidade de simular a imprecisão e ambiguidade do ser humano e porque o problema em questão é de natureza imprecisa.



Figura 9 – Seis graus de maturação do tomate, de acordo com o USDA [3]

Em relação ao problema de classificação de carne, Iqbal *et al.* [56] e Valous *et al.* [57] propuseram classificar presuntos através de imagens. O primeiro analisou 425 amostras de imagens de presunto suíno e 295 de presunto de peru, utilizou os atributos de cor e textura, e três classes. Com o método de Análise Discriminante Linear (Linear Discriminant Analysis - LDA) obteve acurácia de 99,5% para suíno e 90,5% para o peru. Já o segundo, utilizou uma *Multilayer Perceptron Feedforward Adaptive* para classificar 360 imagens de presunto suíno em quatro classes, baseado nos valores RGB, e alcançou 86,1% de acurácia, um resultado considerado bom por ele, devido a representação dos dados de entrada em poucas dimensões.

Outra abordagem para a classificação de carne suína é realizada através de imagens hiperespectrais capturadas com o *Near-infrared Spectroscopy* (NIRS). Seja para classificação com modelos matemáticos [58, 63, 64, 65, 61] ou também com clusterização e RNA [59, 21, 66, 67].

Como por exemplo, Monroy *et al.* [58] utilizaram 240 amostras com valores de cor, textura e exsudação, para classificar entre quatro classes de qualidade de carne suína. Para isso, foi utilizada análise de regressão para selecionar as características mais significativas e análise discriminante para classificação. Com o método de avaliação *leave-one-out* obteve 80% de acurácia e 79% com *cross-validation 5-folds*. Com o uso do método de Análise Discriminante por Mínimos Quadrados Parciais, Prieto *et al.* [65] analisaram 148 amostras de diferentes raças de suínos, classificaram quanto aos dias de *post-mortem*, 2 ou 14 e para determinadas raças e dias alcançaram 100% de acerto. Da mesma forma, Ma *et al.* [61] obtiveram uma média de acerto de 97,73% para classificar 110 amostras entre três classes, com as variáveis de todos os comprimentos de onda e para a fusão dos comprimentos de onda ótimos com as características de textura.

Além da Análise Discriminante Linear, Liu *et al.* [59] utilizaram clusterização k-médias para categorizar a carne de suíno em quatro níveis, utilizando cor, textura e exsudação. Sua maior acurácia foi de 84% para a Análise Discriminante Linear, com um total de 210 combinações de treinamento e teste para obter os resultados de classificação estatística imparcial. Já Chen *et al.* [21] aplicaram também o k-NN e uma RNA com *Backpropagation* em um conjunto de dados com 67 amostras, para determinar o tempo de armazenamento baseado no frescor da amostra. A RNA com *cross-validation* obteve o melhor resultado de 96,20% de acurácia.

Para classificar carne suína entre fresco e congelado-descongelado de acordo com a imagem hiperespectral e textura, Pu *et al.* [67] utilizaram uma Rede Neural Probabilística baseada na classificação Bayesiana. Com um conjunto de dados de 110 amostras, eles realizaram a extração de características de textura de três maneiras diferentes: por histograma, *Gray Level Co-Occurrence Matrix (GLCM)* e *Gray Level-Gradient Co-Occurrence Matrix (GLGCM)*. Para avaliar os resultados obtidos pela RNA, foi utilizada a curva

ROC com acurácia mais alta de 92,02% e comprovou o aumento da acurácia com o uso de textura.

Com o objetivo de determinar o tempo de armazenamento de carne suína e classificar entre 0 dias, 7 dias e 14 dias, Barbon *et al.* [62] utilizaram diversos algoritmos de inteligência computacional e compararam seus resultados. O conjunto de dados era composto por 250 amostras, com valores de pH , *capacidade de retenção de água*, L^* , a^* e b^* . As técnicas aplicadas foram Árvore de Decisão J48, Naïve Bayes, k-NN com $k=3$, *Random Forest*, SVM, MLP e Lógica Fuzzy com duas abordagens diferentes. Para cada técnica foi utilizado o *cross-validation* com 10-*folds* repetidas 30 vezes. Dessa forma, a *Random Forest*, 3-NN, Fuzzy-CHI e Fuzzy-W apresentaram as melhores acurácias, com 94,41%, 93,57%, 93,23% e 92,35%, respectivamente.

Existem diversos trabalhos que classificam a carne suína quanto a sua qualidade, a maioria com o uso de imagens hiperespectrais como visto anteriormente. Porém, apesar do NIRS apresentar vantagens como ser rápido, objetivo, repetível e não destruir a amostra, nem usar reagente [60, 64], ele apresenta barreiras como a alta dimensionalidade dos dados hiperespectrais e pouco intuitivo.

Dessa forma, esse trabalho propõem uma abordagem fuzzy para problemas em que a classificação é realizada por humanos e/ou baseados em métodos tradicionais. Visto que a lógica Fuzzy é amplamente utilizada para problemas de classificação de alimentos quanto sua qualidade, como visto anteriormente para as frutas, ela pode ser uma abordagem mais intuitiva e sem a alta dimensionalidade do NIRS para a classificação de qualidade de carne suína.

5 METODOLOGIA

A metodologia utilizada para execução dos experimentos está ilustrada no diagrama da Figura 10. (1) Experimento 1 - Lógica Clássica: as amostras reais do Dataset Original, foram rotuladas com a lógica Clássica, de acordo com os limiares de cada padrão de avaliação de qualidade. As amostras rotuladas foram separadas em novos datasets, a fim de serem utilizadas como modelo ouro. (2) Experimento 2 - Fuzzy Top-Down: as mesmas amostras não rotuladas utilizadas no experimento anterior, foram classificadas com a abordagem Fuzzy Top-Down com o objetivo de comparar os resultados quanto a quantidade de amostras classificadas e verificar a factibilidade da aplicação do modelo fuzzy. Além de classificar as amostras rotuladas em (1) e calcular a acurácia do modelo para comparar com os resultados do Experimento 3. (3) Experimento 3 - Bottom-Up: os algoritmos de AM foram treinados e testados com os datasets sintéticos, criados a partir dos limiares de cada padrão de avaliação de qualidade. Com os modelos gerados, foi realizada a predição com as bases de amostras classificadas, derivadas de (1), e calculada a acurácia.

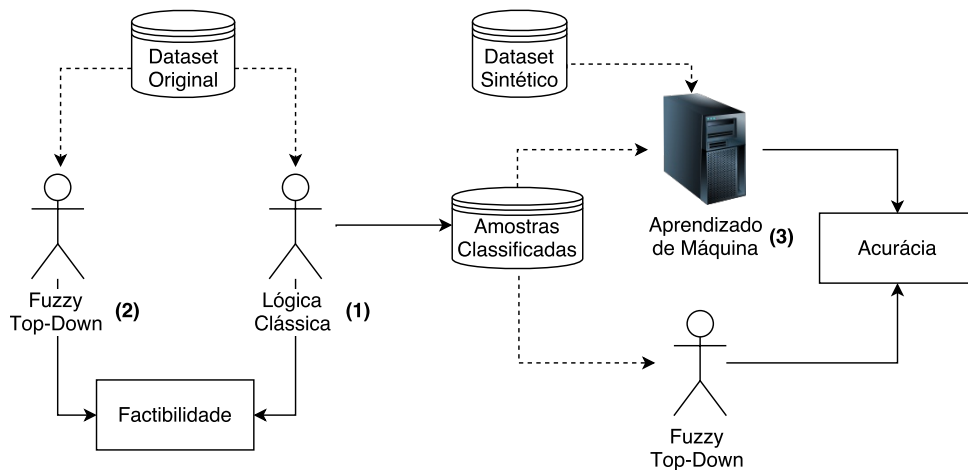


Figura 10 – Diagrama da Metodologia Proposta

Nas próximas seções são descritos os materiais utilizados para os experimentos. Na Seção 5.1 é descrito o problema da classificação de carne suína quanto à qualidade e os padrões utilizados para os experimentos. Na Seção 5.2 são descritos os datasets, tanto a forma de aquisição quanto medidas de dispersão dos atributos. O software utilizado é apresentado na Seção 5.3, seguido da apresentação dos modelos desenvolvidos na Seção 5.4 e na Seção 5.5 uma forma de avaliação de desempenho de algoritmos multidomínio.

5.1 Avaliação de Qualidade de Carne Suína

A preocupação com a qualidade da carne produzida tornou-se cada vez mais importante, tanto quanto a quantidade [18], pois o consumidor tem a qualidade da carne como principal preocupação [19], além de procurar por carnes mais saudáveis. Porém, um grande problema da indústria de carne suína é a falta de uma única definição para qualidade de carne, uma vez que ela é uma combinação de medidas subjetivas e objetivas que variam entre os mercados [20].

A qualidade da carne suína é altamente variável [68] e existem diversas metodologias de classificação [69]. Embora as características de qualidade sejam relacionadas, existem algumas variações independentes dos atributos que levam a uma avaliação imprecisa da qualidade da carne [70, 71]. Por exemplo, os padrões Faucitano [72] e Kauffman [73] apresentam a mesma quantidade de classes, porém o primeiro avalia três atributos, e o segundo dois. O mesmo acontece para Joo [74] e Warner [5].

Além da variedade dos padrões de avaliação da qualidade da carne suína, os intervalos dos atributos foram estabelecidas por conjuntos *crisp*. Em outras palavras, os padrões tradicionais apresentam limites rígidos, o que torna difícil a classificação de amostras que não se encaixam perfeitamente no intervalo esperado do parâmetro, pois elas realizam uma transição abrupta entre uma classe e outra. Por isso, esse problema foi escolhido para aplicação dos modelos fuzzy e realizar a análise de seu desempenho.

Para os experimentos, foram escolhidos quatro padrões da literatura, seguindo o trabalho de diferentes autores. Dois dos quatro padrões avaliam medidas de pH , L^* e CRA , e os outros dois utilizam apenas valores de L^* e CRA , e todos utilizam triplas para representar as classes.

O padrão de Faucitano [72] avalia três atributos e apresenta cinco classes de qualidade, como Kauffman [73], porém, ele avalia apenas L^* e CRA . Joo [74] e Warner [5] classificam em quatro classes, Joo com dois atributos e Warner com os três. A Tabela 4 mostra detalhadamente os intervalos condicionais para cada classe. Pode-se observar que existem algumas lacunas que os padrões não cobrem, como os valores 42, 49 e 50 para CRA do padrão Faucitano, por exemplo.

5.2 Base de Dados

As amostras reais utilizadas nos experimentos foram obtidas de suínos abatidos em frigoríficos comerciais e analisados no Laboratório de Análise de Alimentos e Nutrição Animal (LANA), no departamento de Zootecnia da Universidade Estadual de Londrina. Considerando parâmetros de qualidade, valores de pH , cor (L^*) e Capacidade de Retenção de Água (CRA) foram escolhidos para descrever as amostras.

Tabela 4 – Padrões de Qualidade para Carne suína

Padrão	Classe	pH	L*	CRA
Faucitano	PSE	<6	>50	≥8.1
	PFN	<6	>50	<8.1
	RSE	<6	43-48	≥8.1
	RFN	<6	43-48	<8.1
	DFD	≥6	<42	<4
Joo	PSE	-	≥50	≥6
	RSE	-	≤50	≥6
	RFN	-	≥43	≤6
	DFD	-	≤43	≤6
Kauffman	PSE	-	>58	>5
	RSE	-	52-58	>5
	PFN	-	>58	<5
	RFN	-	52-58	<5
	DFD	-	<52	<5
Warner	PSE	<6	>50	>5
	RSE	<6	42-50	>5
	RFN	<6	42-50	<5
	DFD	≥6	<42	<5

Os valores desses descritores foram obtidos do músculo suíno *longissimus dorsi at lumborum*, removido entre a penúltima e última costela da metade esquerda da carcaça. Para cada amostra, foi medido o *pH* depois de 24 horas do abate e aplicado o medidor de *pH* Testo 205 na carcaça resfriada a $2\pm 2^\circ\text{C}$. A coloração foi avaliada utilizando um colorímetro portátil Minolta® (modelo CR-10 com iluminante C e 8° de inclinação - Tokio, JP) e o componente L^* (*color component*) foi determinado de acordo com o sistema CIEL*a*b*.

O *CRA* foi obtido por duas metodologias: a umidade do filtro de papel (*Filter Paper Wetness* - FPW) [75] e perda de água por gotejamento (*Drip Loss* - DL) [76]. Nesse caso, os valores de *CRA* de algumas amostras foram obtidos por FPW e convertidos para DL para padronizar as unidades de medida. Foi utilizada a fórmula $Y = 0,100707x$ [77] para converter de FPW para DL, em que Y é a capacidade de retenção de água calculada e x é o valor em FPW.

A Tabela 5 mostra algumas informações das sete bases de dados utilizados nesse trabalho. Tal como a quantidade de amostras de cada um e medidas de dispersão, como média, desvio padrão e variância de cada atributo.

Tabela 5 – Quantidade total de amostras e medidas de dispersão para os atributos das bases de dados com amostras reais

Base de Dados	Qtd. Amostras	Medidas de Dispersão	<i>pH</i>	<i>L*</i>	<i>CRA</i>
Dataset 1	73	Média	5,56	52,92	3,34
		Desvio Padrão	0,15	3,35	0,35
		Variância	0,02	11,23	0,12
Dataset 2	306	Média	5,78	53,80	2,88
		Desvio Padrão	0,36	3,21	0,54
		Variância	0,13	10,29	0,29
Dataset 3	405	Média	5,68	52,44	2,71
		Desvio Padrão	0,16	2,62	0,44
		Variância	0,03	6,89	0,19
Dataset 4	117	Média	5,52	50,39	3,17
		Desvio Padrão	0,13	2,56	0,43
		Variância	0,02	6,55	0,18
Dataset 5	54	Média	5,86	50,36	2,94
		Desvio Padrão	0,24	2,31	0,59
		Variância	0,06	5,32	0,35
Dataset 6	44	Média	5,74	55,09	4,99
		Desvio Padrão	0,09	2,83	2,41
		Variância	0,01	8,00	5,82
Dataset 7	48	Média	5,56	56,44	4,97
		Desvio Padrão	0,13	4,63	2,15
		Variância	0,02	21,47	4,64
Dataset 8	144	Média	5,58	54,45	3,30
		Desvio Padrão	0,11	3,35	0,38
		Variância	0,01	11,22	0,15
Dataset 9	70	Média	5,86	52,12	3,62
		Desvio Padrão	0,06	2,33	1,50
		Variância	0,00	5,42	2,26
Dataset 10	40	Média	5,54	52,20	3,45
		Desvio Padrão	0,15	3,66	0,42
		Variância	0,02	13,37	0,18

Por se tratar de procedimentos que envolvem animais, não é possível obter amostras de todas as classes. Determinadas classes são influenciadas pelo manejo do animal no pré-abate [78], pois o nível de estresse durante esse período, influencia na velocidade de

Tabela 6 – Medidas de dispersão para pH, L* e CRA das bases sintéticas geradas para os padrões Faucitano, Joo, Kauffman e Warner

Base de Dados	Medidas de Dispersão	pH	L*	CRA
Faucitano Sintético	Média	4,17	49,22	6,45
	Desvio Padrão	2,96	12,86	4,3
	Variância	8,77	165,33	18,48
Joo Sintético	Média	-	45,54	6,20
	Desvio Padrão	-	14,72	4,19
	Variância	-	216,72	17,53
Kauffman Sintético	Média	-	54,70	5,57
	Desvio Padrão	-	11,22	4,12
	Variância	-	125,85	16,94
Warner Sintético	Média	4,50	45,69	6,02
	Desvio Padrão	3,13	11,36	4,17
	Variância	9,79	129,14	17,36

transformação de glicogênio muscular em ácido lático [79]. Por exemplo, uma carne PSE apresenta alta concentração de ácido lático, ocasionado pelo estresse intenso que o animal sofreu antes do abate e provocou a aceleração da glicólise após o abate [80]. No entanto, se o animal passa por estresse crônico antes do abate, há baixa produção de ácido lático em consequência do esgotamento dos níveis de glicogênio e têm-se uma carne considerada DFD.

Dessa forma, como visto na Seção 2.3, não é possível utilizar os métodos de aprendizado de máquina de forma efetiva. Visto que os datasets com amostras reais estão desbalanceados e não apresentam exemplos de todas as classes, foi necessária a criação de uma base de dados com amostras sintéticas. Os datasets sintéticos são compostos por valores gerados aleatoriamente, dentro dos limites dos atributos de cada padrão. Por isso, para cada padrão foi gerado um dataset com 1000 amostras rotuladas e balanceadas. A Tabela 6 mostra a média, desvio padrão e variância para os atributos dos datasets sintéticos, pH, L* e CRA para Faucitano e Warner, e L* e CRA para Joo e Kauffman.

5.3 R

R¹ é uma linguagem e um ambiente amplamente utilizados, sem custo e *open-source* para computação estatística e gráfica [81]. Pode ser executado em diversas plataformas LINUX, Windows e MacOS. Além disso, ele é facilmente estendido por mais de 5000 pacotes adicionais desenvolvidos pela comunidade R.

¹ <https://www.r-project.org/about.html>

O pacote CARET foi utilizado para a execução dos algoritmos de aprendizado de máquina, pois ele é um conjunto de funções que tentam simplificar o processo de criação de modelos preditivos. Por exemplo, apresenta funções para dividir os conjuntos de dados, realizar o pré-processamento, seleção de atributos, *tunning* do modelo com reamostragem, estimativa da importância da variável. Além de oferecer uma interface uniforme para as funções, como uma forma de padronizar as tarefas comuns, visto que existem diferentes funções de modelagem no R, com diferentes sintaxes para o modelo de treinamento e/ou predição.

Os pacotes dos algoritmos utilizados para classificação estão listados na Tabela 7.

Tabela 7 – Algoritmos utilizados e seus pacotes R

Algoritmo	Método ID	Pacote R
Árvore de Decisão	J48	RWeka
Fuzzy-RBS (Chi)	FRBCS.CHI	frbs
Fuzzy-RBS (Ishibuchi)	FRBCS.W	frbs
k-NN	knn	kknn
<i>Multilayer Perceptron</i>	mlp	RWeka
<i>Naïve Bayes</i>	nb	e1071
<i>Random Forest</i>	rf	randomForest
<i>Support Vector Machine</i>	svmPoly	e1071

Além disso, o pacote FRBS foi utilizado para a construção do modelo fuzzy Top-Down com a função `frbs.gen`. Dessa forma, em vez do algoritmo encontrar os limiares que separam as diferentes classes, essas informações são dadas pelo usuário. Desde a escolha da quantidade de variáveis linguísticas, tipo e quantidade das funções de pertinência, até a construção das regras IF-THEN, não sendo necessário o processo de treinamento [37].

5.4 Modelos

Os modelos de classificação fuzzy foram construídos com funções do pacote FRBS do R. Como visto da Seção 5.3, o pacote oferece duas opções para a geração dos modelos. Uma forma manual, em que o usuário determina a quantidade de variáveis e termos linguísticos, escolhe as formas das funções de pertinência, determina seus intervalos e constrói a base de regras. E outra em que dado um conjunto de dados, o sistema gera as regras de acordo com o que ele aprendeu da base e os intervalos de acordo com a quantidade e forma das variáveis linguísticas escolhidas.

Assim, para a criação do modelo FRBS para a abordagem Top-Down foi utilizada a função `frbs.gen` com as configurações exibidas na Tabela 8. Para o tipo do modelo de inferência foi escolhido o MAMDANI, devido a facilidade da parte consequente das

regras IF-THEN, que é um termo linguístico, ao contrário da opção TSK que requer uma representação de equações lineares em forma de matriz. O método de defuzzificação escolhido foi o da média ponderada (*weighted average method* - WAM), porém o pacote oferece mais outras quatro opções: primeiro dos máximos (*first of maxima* - FIRST.MAX), último dos máximos (*last of maxima* - LAST.MAX), média dos máximos (*mean of maxima* - MEAN.MAX) e o centro de massa modificado (*modified center of gravity* - COG).

Tabela 8 – Parâmetros de configuração para a geração do modelo FRBS com a abordagem Top-Down

Parâmetros de Inferência	Valor
type.model	MAMDANI
type.tnorm	MIN
type.snorm	MAX
type.implication.func	MIN
type.defuz	WAM

Os intervalos e as formas das funções de pertinência foram escolhidas de acordo com os padrões de avaliação de qualidade descritos na Seção 5.1, e podem ser observados na Figura 11.

Os intervalos utilizados para as funções de pertinência de Faucitano, ilustradas na Figura 11a foram: pH.baixo(0, 5, 7), pH.alto(5, 7, 12), L*.baixo(0, 42, 43), L*.medio(42, 43, 48, 50), L*.alto(48, 50, 64), CRA.baixo(0, 4, 5), CRA.medio(0, 7.1, 9.1) e CRA.alto(7.1, 9.1, 10).

Para as funções de pertinência de Joo da Figura 11b, são: L*.xbaixo(0, 42, 44), L*.baixo(0, 49, 51), L*.alto(42, 44, 70), L*.xalto(49, 51, 70), CRA.baixo(0, 5, 7) e CRA.alto(5, 7, 14).

As funções de pertinência de Kauffman da Figura 11c, foram delimitadas como segue: L*.baixo(0, 51, 53), L*.medio(51, 53, 57, 59), L*.alto(57, 59, 66), CRA.baixo(0,4,6) e CRA.alto(4, 6, 14).

Por fim, os intervalos das funções de pertinência para Warner da Figura 11d: pH.baixo(0, 5, 7), pH.alto(5, 7, 12), L*.baixo(0, 41, 43), L*.medio(41, 43, 49, 51), L*.alto(49, 51, 70), CRA.baixo(0, 4, 6) e CRA.alto(4, 6, 10).

Após definir as funções de pertinência, foram criadas as regras IF-THEN para os padrões de acordo com as proposições e condições dos próprios padrões, apenas trocando os intervalos numéricos pelos termos linguísticos. As Tabelas de 9 a 12 exibem as regras para os modelos de Faucitano, Joo, Kauffman e Warner, respectivamente.

Os modelos Bottom-Up foram treinados com o pacote CARET. Para cada padrão de qualidade de carne, o dataset sintético foi particionado com *holdout* estratificado, sendo

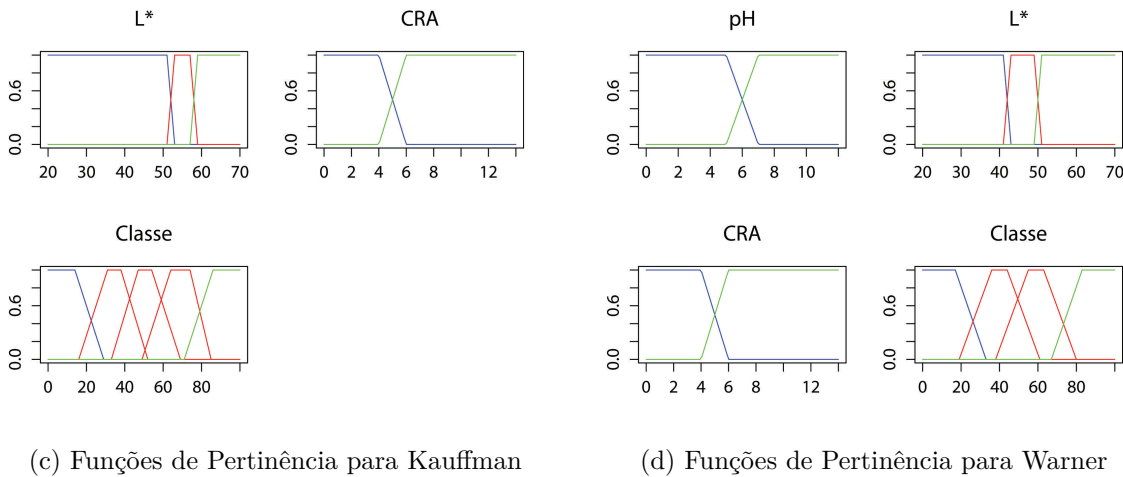
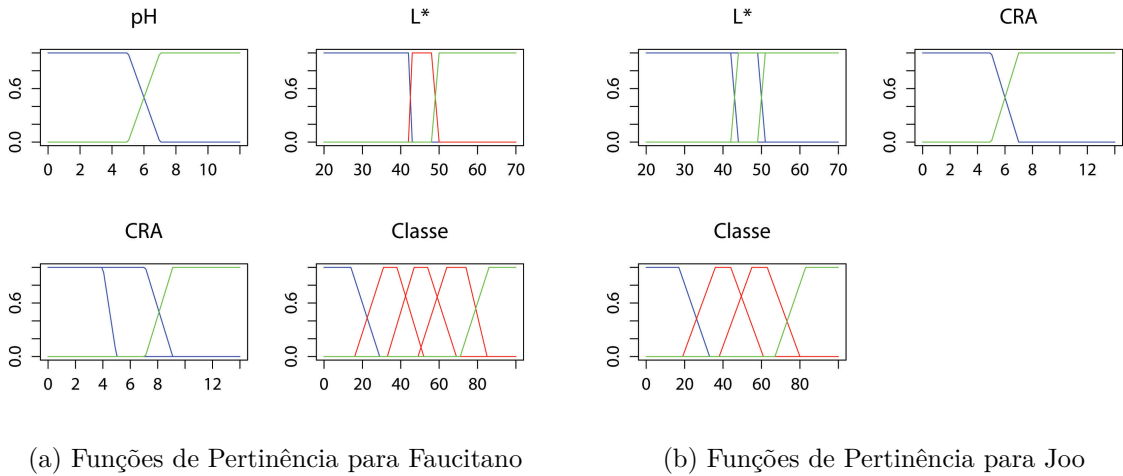


Figura 11 – Funções de Pertinência construídas para cada padrão

Tabela 9 – Regras IF-THEN para o padrão de Faucitano

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16
1	IF	pH	is	baixo	AND	L*	is	alto	AND	cra	is	alto	THEN	Classe	is	PSE
2	IF	pH	is	baixo	AND	L*	is	alto	AND	cra	is	medio	THEN	Classe	is	PFN
3	IF	pH	is	baixo	AND	L*	is	medio	AND	cra	is	alto	THEN	Classe	is	RSE
4	IF	pH	is	baixo	AND	L*	is	medio	AND	cra	is	medio	THEN	Classe	is	RFN
5	IF	pH	is	alto	AND	L*	is	baixo	AND	cra	is	baixo	THEN	Classe	is	DFD

60% para treinamento e os 40% restantes para teste. Todos os algoritmos receberam como entrada os mesmos conjuntos de treinamento e teste. O treinamento foi realizado com *cross-validation* com 10-*folds*, repetidos 10 vezes, com *tuneLength* = 5. E então, com o modelo treinado, foi realizada a predição dos datasets reais, utilizando os rótulos da Lógica Clássica como modelo ouro.

Tabela 10 – Regras IF-THEN para o padrão de Joo

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12
1	IF	L*	is	x.alto	AND	cra	is	alto	THEN	Classe	is	PSE
2	IF	L*	is	baixo	AND	cra	is	alto	THEN	Classe	is	RSE
3	IF	L*	is	alto	AND	cra	is	baixo	THEN	Classe	is	RFN
4	IF	L*	is	x.baixo	AND	cra	is	baixo	THEN	Classe	is	DFD

Tabela 11 – Regras IF-THEN para o padrão de Kauffman

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12
1	IF	L*	is	alto	AND	cra	is	alto	THEN	Classe	is	PSE
2	IF	L*	is	medio	AND	cra	is	alto	THEN	Classe	is	RSE
3	IF	L*	is	alto	AND	cra	is	baixo	THEN	Classe	is	PFN
4	IF	L*	is	medio	AND	cra	is	baixo	THEN	Classe	is	RFN
5	IF	L*	is	baixo	AND	cra	is	baixo	THEN	Classe	is	DFD

Tabela 12 – Regras IF-THEN para o padrão de Warner

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16
1	IF	pH	is	baixo	AND	L*	is	alto	AND	cra	is	alto	THEN	Classe	is	PSE
2	IF	pH	is	baixo	AND	L*	is	medio	AND	cra	is	alto	THEN	Classe	is	RSE
3	IF	pH	is	baixo	AND	L*	is	medio	AND	cra	is	baixo	THEN	Classe	is	RFN
4	IF	pH	is	alto	AND	L*	is	baixo	AND	cra	is	baixo	THEN	Classe	is	DFD

5.5 Avaliação de Desempenho de Modelos Multidomínio

O teste de hipóteses é uma forma de comparar desempenhos de diferentes modelos, além da simples análise de superioridade ou inferioridade de médias. Para comparação de múltiplos domínios recomenda-se o uso do Teste de Friedman [82, 4]. Ele utiliza o valor absoluto da medida de desempenho de cada algoritmo, para cada conjunto de dados, para realizar o ranqueamento (dos melhores para os piores).

Caso a diferença entre os valores médios de posição no ranqueamento de dois algoritmos, for maior ou igual ao valor da diferença crítica (*Critical Difference* - CD), pode-se dizer que eles são estatisticamente diferentes. A CD pode ser calculada com a Equação 5.1, em que A é a quantidade de algoritmos, N é quantidade de conjuntos de dados e q_α pode ser obtido pela estatística Nemenyi [83].

$$CD = q_\alpha \sqrt{\frac{A(A+1)}{6N}} \quad (5.1)$$

Por exemplo, a Figura 12 ilustra a comparação de desempenho de quatro algoritmos, A, B, C e D. O ranqueamento dos algoritmos é realizado de acordo com Friedman e são posicionados em uma escala seguindo os valores médios de suas posições. A CD é colocada acima dessa escala e as barras horizontais conectam modelos que apresentam

similaridades estatísticas.

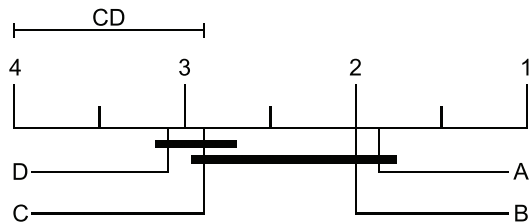


Figura 12 – Exemplo de Avaliação de Desempenho de Múltiplos Modelos, com ranqueamento de Friedman e cálculo da CD com Nemenyi [4]

Assim sendo, de acordo com a Figura 12, pode-se dizer que A foi ligeiramente melhor, B está em média na segunda posição ao avaliar seu desempenho em relação aos conjuntos de dados. O modelo D apresenta um resultado inferior aos modelos A e B, enquanto C não diz nada, pois “um indivíduo não pode pertencer a duas populações diferentes em termos estatísticos” [4].

6 RESULTADOS E DISCUSSÃO

Nesse capítulo são apresentados os resultados dos experimentos para classificação das amostras de carne suína com a lógica Clássica, e as abordagens Fuzzy divididas entre Top-Down e Bottom-Up. Além da análise dos graus de pertinência dos termos linguísticos para as amostras não-classificadas pela lógica Clássica e comparação das abordagens.

6.1 Experimento 1: Lógica Clássica

A classificação com o uso da Lógica Clássica foi realizada por algoritmos com condições se-então aplicados no dataset Original, sendo um algoritmo para cada padrão, respeitando os limiares de cada um. Por exemplo, o Algoritmo 1 para o padrão Warner.

Algorithm 1 Algoritmo para classificação pela Lógica Clássica para o padrão Warner

```

1: for Dataset completo do
2:   if  $pH < 6$  and  $L^* > 50$  and  $CRA > 5$  then
3:      $classe \leftarrow PSE$ 
4:   else
5:     if  $pH < 6$  and  $(L^* \geq 42$  and  $L^* \leq 50)$  and  $CRA > 5$  then
6:        $classe \leftarrow RSE$ 
7:     else
8:       if  $pH < 6$  and  $(L^* \geq 42$  and  $L^* \leq 50)$  and  $CRA < 5$  then
9:          $classe \leftarrow RFN$ 
10:      else
11:        if  $pH \geq 6$  and  $L^* < 42$  and  $CRA < 5$  then
12:           $classe \leftarrow DFD$ 
13:        else
14:           $classe \leftarrow FALSE$ 
15:        end if
16:      end if
17:    end if
18:  end if
19: end for

```

De acordo com os resultados da Tabela 13, o padrão Joo foi capaz de classificar todas as amostras. Isso não quer dizer que esse padrão seja o mais “correto”, mas sim, que os intervalos definidos por ele englobam mais amostras, e que os outros três padrões podem apresentar falhas na definição de seus intervalos ou uma especificação muito rígida para cada parâmetro. Como no padrão do Faucitano que deixa de fora os valores 49 e 50 para o atributo L^* . Em seguida, temos o padrão Kauffman que não atingiu 100% para três datasets. Em terceiro está Faucitano, com porcentagem de classificação entre 70,37% e 97,73%. E por último está Warner, com quantidade de acertos de no máximo 47,92% para um dos datasets com menos amostras.

Tabela 13 – Porcentagem de Amostras Classificadas pela Lógica Clássica

	Total	Faucitano (%)	Joo (%)	Kauffman (%)	Warner (%)
Dataset 1	73	84,93	100,00	100,00	12,33
Dataset 2	306	75,16	100,00	100,00	5,88
Dataset 3	405	85,19	100,00	100,00	16,05
Dataset 4	117	72,65	100,00	100,00	44,44
Dataset 5	54	70,37	100,00	100,00	24,07
Dataset 6	44	97,73	100,00	97,73	31,82
Dataset 7	48	95,83	100,00	97,92	47,92
Dataset 8	144	93,75	100,00	100,00	6,94
Dataset 9	70	87,14	100,00	91,43	30,00
Dataset 10	40	77,50	100,00	100,00	30,00

Os experimentos com abordagem clássica foram aplicados no dataset Original, que são os datasets com amostras reais. As amostras possíveis de serem classificadas foram salvas em datasets separados, um para cada dataset e para cada autor. Os datasets classificados foram utilizados no Experimento 3, como modelo ouro, para o cálculo da acurácia.

6.2 Experimento 2: Top-Down

A abordagem Top-Down visa dois objetivos: verificar a viabilidade do método de classificação fuzzy quando aplicado no Dataset Original e analisar a acurácia quando aplicado nos datasets classificados. Os modelos FRBS para essa abordagem foram construídos de forma manual com a função `frbs.gen` do pacote, ou seja, os intervalos das funções de pertinência e a base de regras foram construídas baseadas nas informações dos padrões.

O modelo do Faucitano foi construído com três variáveis linguísticas: pH , L^* e CRA ; dois termos linguísticos para pH e três para L^* e CRA ; cinco saídas: PSE, PFN, RSE, RFN, DFD; e cinco regras. Para o modelo do Joo, foram utilizadas duas variáveis linguísticas, pois ele avalia apenas valores de L^* e CRA ; dois termos linguísticos para cada variável; e quatro saídas: PSE, RSE, RFN, DFD; e quatro regras.

Kauffman também avalia apenas L^* e CRA , por isso foram criadas duas variáveis linguísticas; três termos linguísticos para L^* e dois para CRA ; cinco classes de saída: PSE, RSE, PFN, RFN, DFD; e cinco regras. Por fim, o modelo do Warner tem três variáveis linguísticas, pH , L^* e CRA ; dois termos linguísticos para pH e CRA e três para L^* ; quatro saídas: PSE, RSE, RFN, DFD; e quatro regras.

Dessa forma, a Tabela 14 mostra a porcentagem de amostras classificadas com a

abordagem Top-Down para cada dataset e autor. Novamente, padrão de Joo foi capaz de classificar todas as amostras. Ao contrário do experimento anterior, Faucitano ficou em segundo lugar com a porcentagem mais baixa de 98,63%, seguido de Kauffman com valor mínimo de 97,73%, e por fim, Warner com valores entre 18,30% e 68,75%.

Tabela 14 – Porcentagem de Amostras Classificadas pela Abordagem Top-Down

	Total	Faucitano (%)	Joo (%)	Kauffman (%)	Warner (%)
Dataset 1	73	98,63	100,00	100,00	24,66
Dataset 2	306	99,02	100,00	100,00	18,30
Dataset 3	405	100,00	100,00	100,00	29,14
Dataset 4	117	100,00	100,00	100,00	59,83
Dataset 5	54	100,00	100,00	100,00	55,56
Dataset 6	44	100,00	100,00	97,73	65,91
Dataset 7	48	100,00	100,00	100,00	68,75
Dataset 8	144	100,00	100,00	100,00	18,75
Dataset 9	70	100,00	100,00	98,57	57,14
Dataset 10	40	100,00	100,00	100,00	40,00

Para avaliar a acurácia do modelo, foi realizada a predição dos datasets classificados pela lógica Clássica de acordo com cada padrão. Como a quantidade de amostras classificadas pelo Lógica Clássica foi diferente para cada padrão, o total de amostras pode ser visto do lado da acurácia na Tabela 15.

Tabela 15 – Acurácia da Classificação Abordagem Top-Down (Acc e total de amostras)

	Faucitano (%)	Joo (%)	Kauffman (%)	Warner (%)
Dataset 1	100,00 (62)	100,00 (73)	97,26 (73)	100,00 (9)
Dataset 2	100,00 (230)	100,00 (306)	92,48 (306)	100,00 (18)
Dataset 3	99,71 (345)	100,00 (405)	98,52 (405)	100,00 (65)
Dataset 4	100,00 (85)	100,00 (117)	99,15 (117)	100,00 (52)
Dataset 5	100,00 (38)	100,00 (54)	98,15 (54)	100,00 (13)
Dataset 6	72,09 (43)	93,18 (44)	65,12 (43)	100,00 (14)
Dataset 7	100,00 (46)	95,83 (48)	78,72 (47)	100,00 (23)
Dataset 8	100,00 (135)	100,00 (144)	84,03 (144)	100,00 (10)
Dataset 9	83,61 (61)	97,14 (70)	90,62 (64)	100,00 (21)
Dataset 10	100,00 (31)	100,00 (40)	92,50 (40)	100,00 (12)

Como pode ser visto, Joo foi capaz de predizer os datasets completos com a acurácia mais baixa de 93,18% para o Dataset 6. Em seguida, está o padrão Faucitano, com menor acurácia de 72,09% também para o Dataset 6, mas ele conseguiu um resultado

melhor para o Dataset 7 em comparação com Joo e Kauffman, com diferença de 1 ou 2 amostras a menos.

Apesar de apresentar datasets com quantidades parecidas com o Joo, Kauffman ficou para trás, com o pior resultado de 65,12% e melhor de 99,15%. Warner obteve acurácia de 100% para todos, porém é importante ressaltar que os datasets preditos por ele tinham o tamanho bastante reduzido em comparação aos outros, visto que a Lógica Clássica para Warner foi capaz de classificar de 5,88% a no máximo 47,92% dos Datasets Originais, como visto na Seção 6.1.

6.3 Experimento 3: Bottom-Up

Algoritmos de classificação de aprendizado supervisionado requerem datasets rotulados, com exemplos de todas as classes do problema. Como os datasets reais utilizados nesse trabalho não atendem a esse requisito, foi necessária a criação e utilização de datasets sintéticos para cada padrão de avaliação de qualidade. Os datasets sintéticos foram utilizados para o treinamento e teste dos modelos. Então, o modelo treinado foi aplicado aos datasets reais para predição.

A acurácia dos modelos foi calculada a partir da matriz de confusão, tomando os resultados da classificação pela Lógica Clássica como modelo ouro. Os resultados podem ser observados nas Figuras 13 e 14.

Em geral, ao realizar a análise por datasets, pode-se concluir que a RF apresenta melhores resultados para todos os padrões. Enquanto a MLP apresenta as acurácias mais baixas, com valores iguais ou próximos a 0 para três dos quatro padrões.

Em relação ao comportamento dos padrões, Joo alcançou as acurácias mais altas para a maioria dos seus datasets. Seguido por Warner, lembrando que seus datasets têm a quantidade de amostras reduzidas. Por último, está o padrão Faucitano com valores altos apenas para a RF, ao contrário dos outros experimentos, em que ele normalmente encontrava-se em segundo lugar.

Faucitano se mostrou o padrão mais instável e com acurácias mais baixas. Os resultados para o Dataset 5, por exemplo, variam de 15,79% a 42,11%, exceto pela RF. Enquanto para o Dataset 6 alcança a acurácia mais alta de 88,37% e a mais baixa de 13,95% e o Dataset 7 apresenta valores mais parecidos, que variam de 36,96% a 63,04% ao considerar a RF. Isso pode ter ocorrido, pois, além da quantidade de amostras de teste reduzida dos datasets, o padrão Faucitano avalia um parâmetro (pH) a mais do que Joo e Kauffman, tornando-o mais restritivo, e portanto, menos amostras são classificadas.

Para melhor comparação entre os métodos de AM, foram construídos gráficos de CD para os quatro padrões de avaliação de qualidade, a partir do ranqueamento de

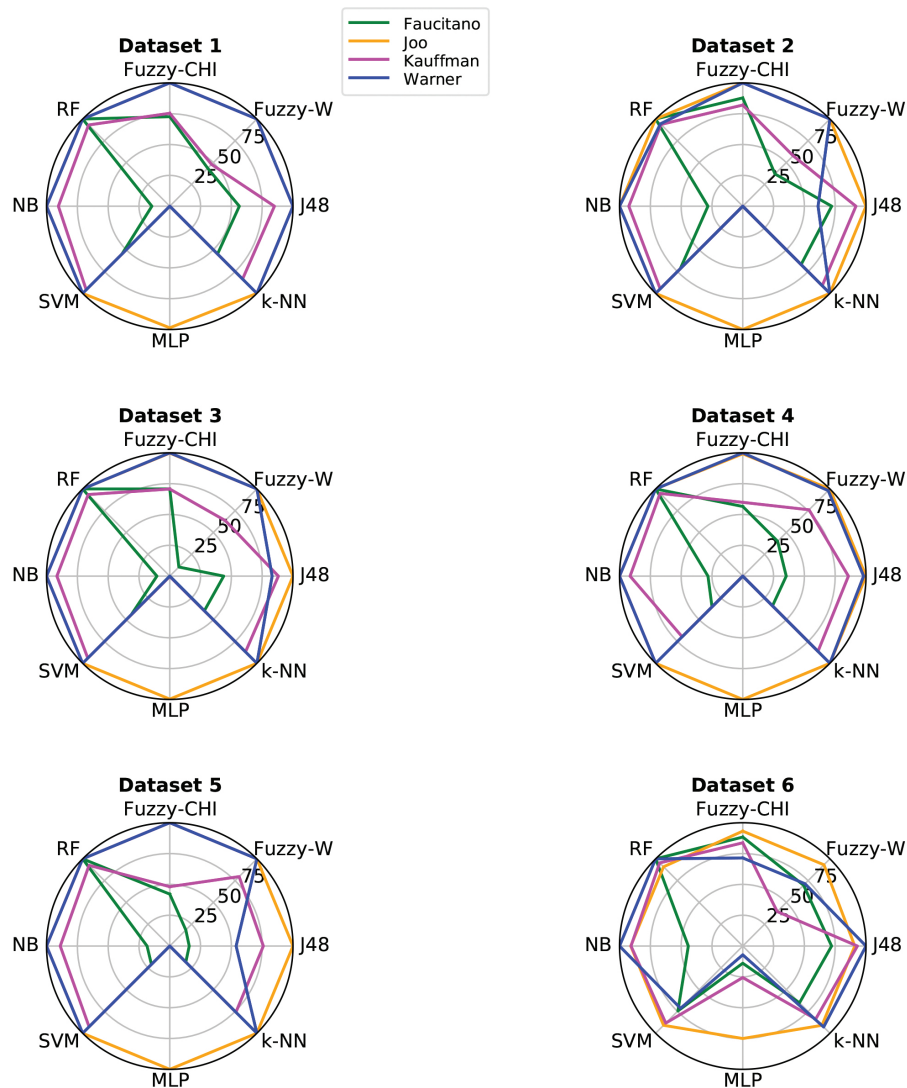


Figura 13 – Acurácia para os Datasets 1 a 6, em relação aos algoritmos de aprendizado de máquina (Fuzzy-CHI, Fuzzy-W, J48, k-NN, MLP, SVM, NB e RF) e padrões de avaliação de carne (Faucitano, Joo, Kauffman e Warner)

Friedman e o cálculo do Nemenyi. Para todos, o valor de CD encontrado foi igual a 3,32 e $\alpha = 0,05$.

Desse modo, de acordo com a Figura 15, a RF apresentou-se em média na primeira posição, conforme o ranqueamento de Friedman, e a MLP em último, como era esperado. Nenhuma conclusão pode ser tomada dos demais algoritmos, pois eles pertencem a dois grupos diferentes estatisticamente. Na Figura 16 pode-se observar que para Joo todos os algoritmos são estatisticamente similares, o Fuzzy-W está no ranking de Friedman mais alto comparado aos outros, mas não apresenta grande destaque.

De acordo com a Figura 17, a RF e a SVM encontram-se nos rankings mais altos, em relação aos outros algoritmos. Enquanto o Fuzzy-CHI, Fuzzy-W e a MLP aparecem

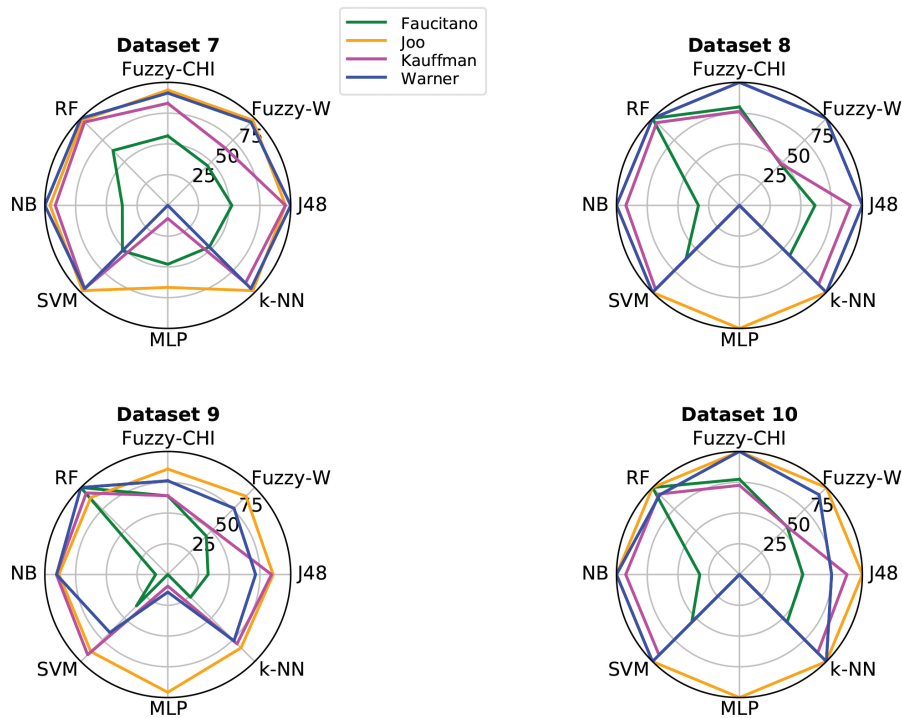


Figura 14 – Acurácia para os Datasets 7 a 10, em relação aos algoritmos de aprendizado de máquina (Fuzzy-CHI, Fuzzy-W, J48, k-NN, MLP, SVM, NB e RF) e padrões de avaliação de carne (Faucitano, Joo, Kauffman e Warner)

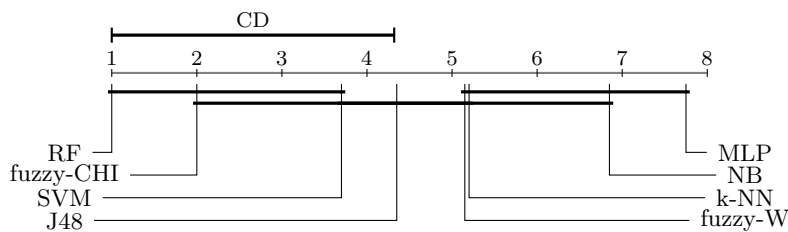


Figura 15 – Análise de Desempenho dos Algoritmos de Aprendizado de Máquina (Fuzzy-CHI, Fuzzy-W, J48, k-NN, MLP, SVM, NB e RF) por Dataset para o Padrão Faucitano, com $CD = 3,32$ e $\alpha = 0,05$

nos últimos lugares, em que as abordagens Fuzzy ocupam as posições entre 6 e 7, e a MLP na posição 8. O caso do Warner, Figura 18, é parecido com o padrão Joo, em que a maioria dos algoritmos são similares e apresentam rankings próximos, com a maioria entre 3 e 5. Com a melhor posição está a NB, em último a MLP e o Fuzzy-W ocupando duas populações diferentes, estatisticamente.

Conforme os resultados vistos, pode-se dizer que nenhum algoritmo é unânime para classificação de amostras para todos os padrões. Pois mesmo com resultados altos da RF para a maioria dos datasets, pela comparação realizada pelo cálculo da CD, ela fica com o primeiro lugar ou próximo para Faucitano e Kauffman. Na análise para o padrão Joo, a RF se encontra em quarto lugar no ranking, atrás do Fuzzy-W e junto com a SVM.

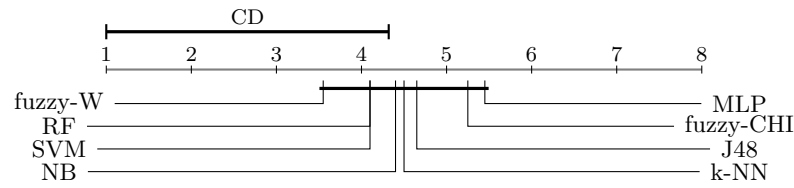


Figura 16 – Análise de Desempenho dos Algoritmos de Aprendizagem de Máquina (Fuzzy-CHI, Fuzzy-W, J48, k-NN, MLP, SVM, NB e RF) por Dataset para o Padrão Joo, com $CD = 3,32$ e $\alpha = 0,05$

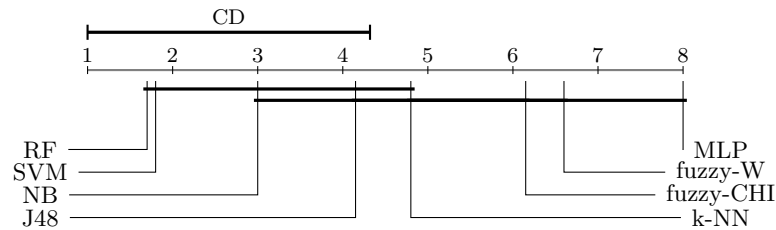


Figura 17 – Análise de Desempenho dos Algoritmos de Aprendizagem de Máquina (Fuzzy-CHI, Fuzzy-W, J48, k-NN, MLP, SVM, NB e RF) por Dataset para o Padrão Kauffman, com $CD = 3,32$ e $\alpha = 0,05$

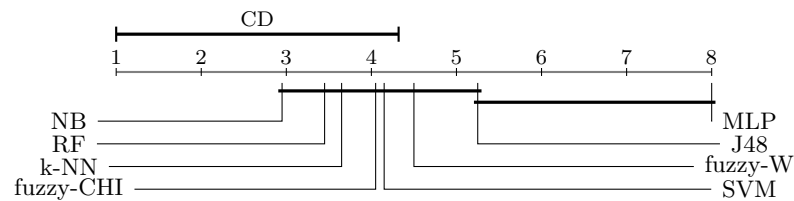


Figura 18 – Análise de Desempenho dos Algoritmos de Aprendizagem de Máquina (Fuzzy-CHI, Fuzzy-W, J48, k-NN, MLP, SVM, NB e RF) por Dataset para o Padrão Warner, com $CD = 3,32$ e $\alpha = 0,05$

Já em relação do padrão Warner, a RF está perto do terceiro lugar, atrás da NB e perto do k-NN.

6.4 Análise dos Graus de Pertinência das Amostras Não-Factíveis

Além dos experimentos das Sessões 6.1 a 6.3, uma última análise foi realizada com a abordagem Top-Down e as amostras não classificadas pela lógica Clássica. Como dito anteriormente, se o modelo for construído manualmente, é possível acessar os graus de pertinência de cada atributo e de cada regra. Dessa forma, foram verificados os graus de pertinência dos atributos das amostras não-classificadas para investigar o porquê não foi possível classificá-las.

As Figuras 19 a 21 ilustram os graus de pertinência de cada termo linguístico, de acordo com cada padrão, para as amostras não classificadas com a lógica Clássica. No eixo x estão os termos linguísticos para o padrão, no eixo y estão as amostras analisadas

e as cores variam de vermelho para amarelo, verde e azul, sendo vermelho os termos linguísticos com grau de pertinência igual a 0, e azul igual a 1.

A classificação de acordo com o padrão de Faucitano obteve 196 amostras não-classificadas, 7 para Kauffman e 934 para Warner, como o padrão de Joo foi capaz de classificar todas as amostras, ele não foi utilizado nessa análise. Esses gráficos têm o objetivo de demonstrar a influência de cada atributo para a classificação. Por exemplo, pode-se perceber que para Faucitano e Warner, valores de pH não são decisivos na escolha, pois a maioria apresenta graus intermediários, o que deixa a tomada de decisão para os outros atributos.

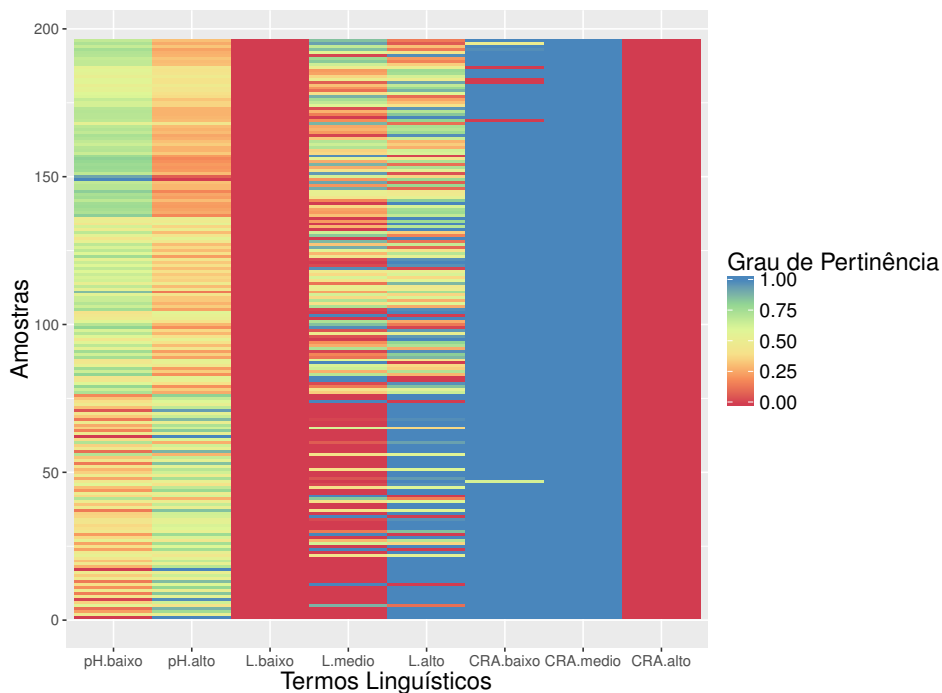


Figura 19 – Amostras não-factíveis para Faucitano

Como exemplo, tem-se a amostra 0 para o padrão de Faucitano, na Figura 19 que apresenta valores de: $pH=7,1$, $L^*=50,0333$ e $CRA=2,9926$. Ela obteve valores 1 para pH.alto, L*.alto, CRA.baixo e CRA.média e zero para os outros termos linguísticos. Porém, não existe nenhuma regra igual a *IF* $pH = alto$ *AND* $L^* = alto$ e $CRA = baixo$ ou *IF* $pH = alto$ *AND* $L^* = alto$ e $CRA = médio$, por isso ela não pôde ser classificada de acordo com a lógica Clássica.

As amostras não-factíveis de Kauffman, na Figura 20, apresentaram todas o mesmo comportamento. Obtiveram graus de pertinência altos para L*.baixo e CRA.alto, e baixos para os outros atributos, porém não existe uma regra na qual esse padrão se encaixe. Além disso, de acordo com a classificação fuzzy Top-Down, a amostra 6 por exemplo, com valores de pertinência de $L^*.baixo=0,9$, $L^*.medio=0,1$, $L^*.alto=0$, $CRA.baixo=0,3$ e $CRA.alto=0,7$, foi atribuída à classe DFD, pois o maior grau de pertinência é de L*.baixo,

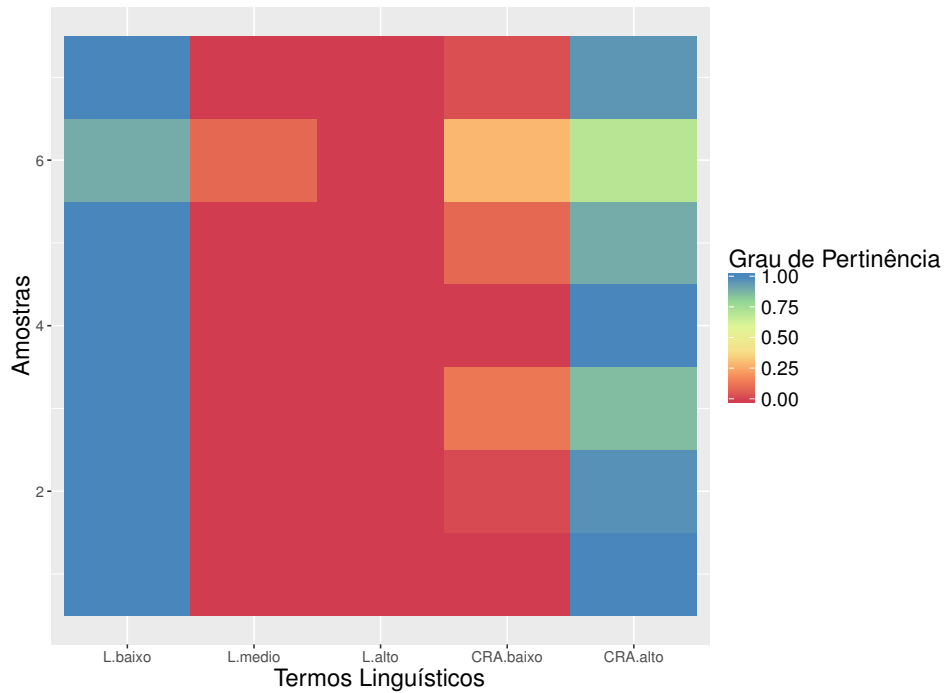


Figura 20 – Amostras não-factíveis para Kauffman

e de acordo com as regras, a única classe que tem $L^*.baixo$ é a DFD.

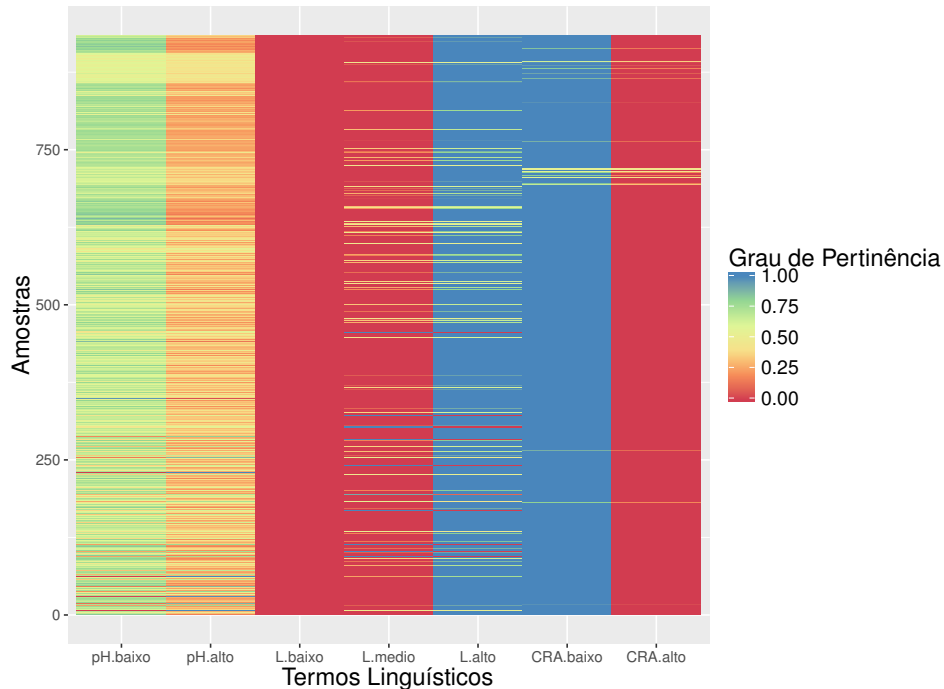


Figura 21 – Amostras não-factíveis para Warner

Warner é o padrão que mais apresenta valores intermediários para pH . Ao verificar os graus de pertinência de duas amostras, pode-se observar dois comportamentos diferentes. A amostra 887 apresenta graus de pertinência de: $pH.baixo=0,57$, $pH.alto=0,43$, $L^*.baixo=0$, $L^*.medio=0$, $L^*.alto=1$, $CRA.baixo=1$, $CRA.alto=0$, que poderia ser assimi-

lada com a regra para PSE pela lógica Clássica, se não fosse o valor de CRA. A segunda amostra é a 888 com graus de pertinência de $pH.baixo=0,54$, $pH.alto=0,46$, $L*.baixo=0$, $L*.medio=0,07$, $L*.alto=0,93$, $CRA.baixo=1$, $CRA.alto=0$ um caso bem parecido, e que não se encaixa nas regras de acordo com a lógica Clássica.

Porém, de acordo com a classificação fuzzy, a primeira amostra recebeu o rótulo de PSE enquanto a segunda, foi rotulada como RFN. Provavelmente, isso pode ter ocorrido pelo fato de existir um pequeno grau de pertinência da amostra 888 para $L*.medio$, fazendo com que a similaridade da amostra tenda mais para Regra 4 do modelo. Nesse caso, mesmo com grau de pertinência máximo para CRA , foi o atributo L^* que mais influenciou na decisão.

6.5 Comparação Lógica Clássica vs. Top-Down

As Seções 6.1, 6.2 e 6.3 descreveram os experimentos com as três abordagens de classificação aplicadas aos datasets originais, classificados e sintéticos para quatro diferentes padrões de classificação. A Figura 22 ilustra a comparação dos resultados da classificação seguindo a lógica Clássica e a abordagem Top-Down.

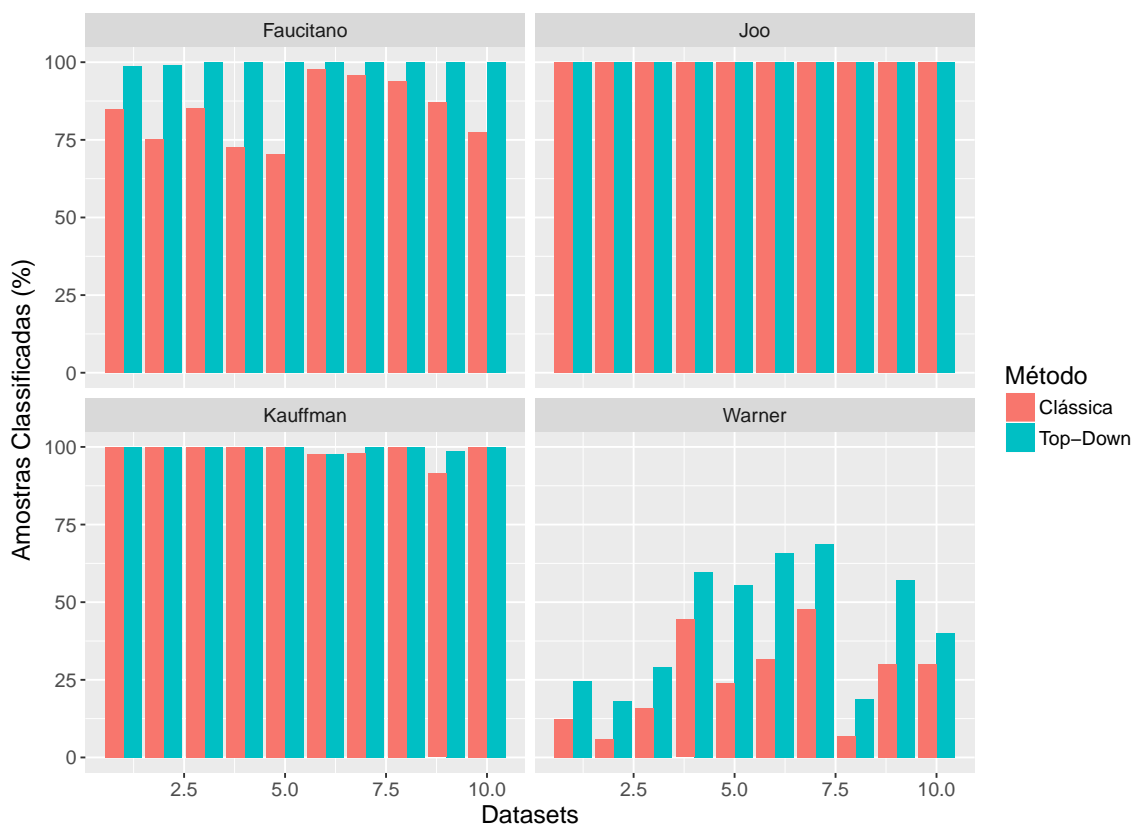


Figura 22 – Lógica Clássica vs. Top-Down

Como pode ser observado, a abordagem Top-Down se igualou e até mesmo superou a quantidade de amostras classificadas pela lógica Clássica. Para o padrão de Joo

ambas classificaram 100% das amostras, e para Kauffman a Clássica se manteve acima de 91,43% e a Top-Down 97,73%. Para o padrão de Faucitano, o número de amostras classificadas obteve resultados expressivos quando utilizada a segunda abordagem, pois diversos datasets aumentaram para 100%, com menor resultado de 98,63%, enquanto a lógica Clássica apresentou maior resultado de 97,73% e 70,37% como menor. No entanto, o padrão Warner é o que apresenta mais problemas, pois seus resultados do Experimento 1 variam de 5,88% para o Dataset 2 e 47,92% para o Dataset 7; e entre 18,30% para o Dataset 2 e 68,75% para o Dataset 7.

6.6 Comparação Top-Down vs. Bottom-Up

Da mesma forma que a Seção 6.5 a Tabela 16 exibe a comparação entre a abordagem Fuzzy Top-Down em relação aos algoritmos de aprendizado de máquina. São apresentados em negrito os resultados do Fuzzy Top-Down e o algoritmo que ficou no ranking mais alto, de acordo com a estatística de Nemenyi.

De acordo com a Tabela 16, a acurácia do Fuzzy Top-Down se aproxima ou mesmo supera os resultados dos outros algoritmos, inclusive o considerado como maior ranking no Nemenyi. Dessa forma, como não há um algoritmo unânime para os quatro padrões, o uso do Fuzzy Top-Down torna-se ainda mais interessante. Pois, além de alcançar bons resultados, o modelo não depende do processo de treinamento, só precisa construir um modelo para cada padrão, sem a necessidade da escolha do algoritmo.

De acordo com os experimentos, a lógica Clássica se mostrou insuficiente para o problema de classificação de carne suína quanto a sua qualidade. Pois, os padrões muitas vezes são criados com limites *crisp*, rígidos, que pode deixar de fora amostras com parâmetros que se encontram em diferentes classes.

Alguns pontos de comparação entre as duas abordagens Fuzzy, Top-Down e Bottom-Up de acordo com suas vantagens e desvantagens: (1) A abordagem Top-Down apresenta a vantagem de análise e fácil interpretação do modelo e dos resultados, pelo fato de ser possível acessar os graus de pertinência dos atributos e das regras, a fim de analisar o porquê de determinadas decisões do sistema, enquanto a Bottom-Up se torna uma caixa preta; (2) Quanto a dependência de um especialista, a segunda abordagem leva vantagem, pois a construção dos modelos é mais fácil, principalmente com o uso do CARET; (3) Por se tratar de uma abordagem de aprendizado supervisionado, a Bottom-Up apresenta restrições quanto ao dataset, como ele ser rotulado e balanceado para obter melhores resultados, enquanto a Top-Down não possui essas restrições, pois o modelo é construído a partir das entradas do usuário; (4) Ao se tratar de tempo para construção do modelo, a Fuzzy Top-Down leva vantagem, pois no caso da abordagem Bottom-Up o tempo irá depender do algoritmo de AM e dos parâmetros de *tunning* e *cross-validation* escolhidos,

Tabela 16 – Comparação das acurácias da classificação dos Datasets Originais, entre o Fuzzy Top-Down e os oito algoritmos de aprendizado de máquina

	Alg.	D ₁ (%)	D ₂ (%)	D ₃ (%)	D ₄ (%)	D ₅ (%)	D ₆ (%)	D ₇ (%)	D ₈ (%)	D ₉ (%)	D ₁₀ (%)
Faucitano	TopDown	100,00	100,00	99,71	100,00	100,00	72,09	100,00	100,00	83,61	100,00
	FuzzyW	43,55	36,96	10,43	40,00	18,42	69,77	45,65	47,41	44,26	54,84
	FuzzyChi	72,58	87,83	70,72	56,47	42,11	88,37	56,52	80,00	63,93	77,42
	J48	56,45	72,17	43,77	35,29	15,79	72,09	52,17	61,48	32,79	51,61
	k-NN	54,84	66,96	39,71	34,12	18,42	65,12	47,83	57,78	26,23	54,84
	MLP	0	0	0	0	0	13,95	47,83	0	0	0
	NB	14,52	28,26	10,43	28,24	18,42	44,19	36,96	33,33	9,84	32,26
	RF	100,00	100,00	100,00	100,00	100,00	100,00	63,04	100,00	100,00	100,00
	SVM	54,84	71,74	44,06	35,29	21,5	74,42	52,17	61,48	36,07	54,84
	Joo	TopDown	100,00	100,00	100,00	100,00	100,00	93,18	95,83	100,00	97,14
FuzzyW		100,00	100,00	100,00	100,00	100,00	93,18	97,92	100,00	90,00	100,00
FuzzyChi		100,00	100,00	99,75	99,15	100,00	93,18	93,75	100,00	85,71	100,00
J48		100,00	100,00	100,00	100,00	100,00	90,91	95,83	100,00	85,71	100,00
k-NN		100,00	100,00	100,00	100,00	100,00	90,91	97,92	100,00	84,29	100,00
MLP		98,63	100,00	99,75	100,00	100,00	75,00	66,67	100,00	95,71	100,00
NB		100,00	100,00	100,00	100,00	100,00	90,91	95,83	100,00	88,57	100,00
RF		100,00	100,00	100,00	100,00	100,00	90,91	97,92	100,00	88,57	100,00
SVM		100,00	100,00	100,00	100,00	100,00	90,91	97,92	100,00	88,57	100,00
Kauffman		TopDown	97,26	92,48	98,52	99,15	98,15	65,12	78,72	84,03	90,62
	FuzzyW	47,95	57,84	63,70	76,07	79,63	39,53	65,96	47,92	51,56	55,00
	FuzzyChi	75,34	82,03	70,62	59,83	48,15	83,72	82,98	76,39	64,06	72,50
	J48	84,93	91,83	88,15	85,47	75,93	93,02	93,62	90,28	84,38	87,50
	k-NN	83,56	91,18	86,91	86,32	75,93	83,72	89,36	90,97	79,69	90,00
	MLP	0	0	0,25	0	0	25,58	10,64	0	9,38	0
	NB	90,41	92,48	91,60	91,45	88,89	90,70	91,49	92,36	90,62	92,50
	RF	93,15	93,79	92,58	94,87	92,59	95,35	95,74	95,14	93,75	92,50
	SVM	95,89	94,77	94,07	96,58	92,59	88,37	95,74	96,53	92,19	92,50
	Joo	TopDown	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
FuzzyW		100,00	100,00	100,00	98,08	100,00	71,43	95,65	100,00	76,19	91,67
FuzzyChi		100,00	100,00	100,00	100,00	100,00	71,43	91,30	100,00	76,19	100,00
J48		100,00	61,11	83,08	98,08	53,85	100,00	100,00	100,00	71,43	75,00
k-NN		100,00	100,00	100,00	100,00	100,00	92,86	95,65	100,00	76,19	100,00
MLP		0	0	0	0	0	7,14	0	0	14,29	0
NB		100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	90,48	100,00
RF		100,00	94,44	100,00	100,00	100,00	100,00	100,00	100,00	100,00	91,67
SVM		100,00	100,00	100,00	100,00	100,00	71,43	95,65	100,00	66,67	100,00

por exemplo, o tempo de treinamento e teste para o Fuzzy-W é muito maior do que para o k-NN.

7 CONCLUSÃO

Devido a rigidez dos limites da lógica Clássica, muitas vezes ela não é capaz de lidar com todas as amostras dado um dataset com amostras reais, como foi comprovado na Seção 6.1. Dessa forma, o uso da lógica Fuzzy com seus graus de pertinência torna-se mais apropriado quando as amostras muitas vezes não tem todos os valores dos atributos pertencentes a mesma classe ou próximos aos limiares.

Quando comparada com a lógica Clássica, a abordagem Fuzzy Top-Down obteve melhores resultados, sendo capaz de igualar, ou até mesmo superar, a quantidade de amostras classificadas. Além de apresentar acurácias próximas a 100% para a maioria dos datasets e padrões, considerando apenas as amostras rotuladas. Enquanto os algoritmos de aprendizado de máquina apresentaram resultados diversos, como o Fuzzy-W que hora alcançou altas acurácias para Joo e Warner, mas baixas para Faucitano e Kauffman, ou a MLP com resultados próximos a 0 para a maioria dos datasets e a RF com a prevalência de acurácias próximas a 100%.

Ambas as abordagens fuzzy Top-Down e Bottom-Up resolvem o problema das amostras não-factíveis para a lógica Clássica, cada uma com suas particularidades. Como a Bottom-Up que precisa de um dataset previamente rotulado, sendo necessário a construção de um dataset com amostras sintéticas no processo de treinamento, pois o dataset real não tinha exemplos de todas as classes. A abordagem Top-Down é uma alternativa que atenua esse requisito, porém demanda mais detalhamento para construção do modelo.

Portanto, de acordo com os resultados obtidos a partir dos experimentos, pode-se comprovar a viabilidade e desempenho da utilização de Sistemas de Inferência Fuzzy para classificação de amostras não-factíveis para a lógica Clássica, com mais vantagens ao utilizar a abordagem fuzzy Top-Down. Pois, alcançou altas acurácias para todos os datasets, com valores próximos aos obtidos pelos algoritmos de aprendizado de máquina de melhor desempenho, é um modelo descritivo de fácil entendimento e análise dos resultados. Além do fato de nenhum algoritmo de aprendizado de máquina ser bom de forma unânime para todos os padrões.

Visto que existem diversos padrões para avaliação da qualidade de carne suína, como trabalhos futuros são propostas duas abordagens, uma seria a fusão de padrões de avaliação de qualidade de carne em um único modelo fuzzy, com o objetivo de facilitar a usabilidade do usuário, que não precisará escolher um padrão de avaliação e terá como resposta a qualidade da carne. Outra abordagem seria o uso de meta-learning para indicar qual padrão de avaliação é melhor utilizar de acordo com a análise da base de dados.

REFERÊNCIAS

- [1] DEAN, J. *Big data, data mining, and machine learning: value creation for business leaders and practitioners*. Hoboken, New Jersey: John Wiley & Sons, 2014.
- [2] MENDEL, J. M. Fuzzy logic systems for engineering: a tutorial. *Proceedings of the IEEE*, Ieee, v. 83, n. 3, p. 345–377, March 1995.
- [3] GOEL, N.; SEHGAL, P. Fuzzy classification of pre-harvest tomatoes for ripeness estimation—an approach based on automatic rule learning using decision tree. *Applied Soft Computing*, Elsevier, v. 36, p. 45–56, 2015.
- [4] BRAGA, A. de P.; CARVALHO, A. C. P. de Leon Ferreira de; LUDERMIR, T. B. *Redes neurais artificiais: teoria e aplicações*. Brasil: LTC Editora, 2007.
- [5] WARNER, R. D.; KAUFFMAN, R. G.; GREASER, M. L. Muscle protein changes post mortem in relation to pork quality traits. *Meat Science*, v. 45, n. 3, p. 339–352, 1997.
- [6] COPPIN, B. *Artificial intelligence illuminated*. USA: Jones & Bartlett Learning, 2004.
- [7] ZADEH, L. A. Fuzzy sets. *Information and control*, Elsevier, v. 8, n. 3, p. 338–353, 1965.
- [8] GUYON, I. et al. *Feature extraction: foundations and applications*. Heidelberg, Germany: Springer, 2008. v. 207.
- [9] TALON, A. F.; MADEIRA, E. R. M.; TOLEDO, M. B. F. Self-adaptive fuzzy architecture to predict and decrease e-contract violations. In: *2014 Brazilian Conference on Intelligent Systems*. São Carlos, SP, Brasil: [s.n.], 2014. p. 294–299.
- [10] LIU, K. et al. Fuzzy classification of young women’s lower body based on anthropometric measurement. *International Journal of Industrial Ergonomics*, Elsevier, v. 55, p. 60–68, 2016.
- [11] LUPO, T. A fuzzy framework to evaluate service quality in the healthcare industry: An empirical case of public hospital service evaluation in sicily. *Applied Soft Computing*, Elsevier, v. 40, p. 468–478, 2016.
- [12] ABIYEV, R. H. et al. Assessment of food security risk level using type 2 fuzzy system. *Procedia Computer Science*, Elsevier, v. 102, p. 547–554, 2016.
- [13] UPADHYAY, R.; SEHWAG, S.; MISHRA, H. N. Electronic nose guided determination of frying disposal time of sunflower oil using fuzzy logic analysis. *Food Chemistry*, Elsevier, v. 221, p. 379–385, 2017.
- [14] KAVDIR, I.; GUYER, D. E. Apple grading using fuzzy logic. *Turkish Journal of Agriculture and Forestry*, The Scientific and Technological Research Council of Turkey, v. 27, n. 6, p. 375–382, 2004.

- [15] MUSTAFA, N. B. A. et al. Agricultural produce sorting and grading using support vector machines and fuzzy logic. In: *2009 IEEE International Conference on Signal and Image Processing Applications*. Kuala Lumpur, Malaysia: [s.n.], 2009. p. 391–396.
- [16] ALAVI, N. Quality determination of mozafati dates using mamdani fuzzy inference system. *Journal of the Saudi Society of Agricultural Sciences*, Elsevier, v. 12, n. 2, p. 137–142, 2013.
- [17] OTHMAN, M. et al. Fuzzy ripening mango index using rgb colour sensor model. *Researchers World*, Educational Research Multimedia & Publications, v. 5, n. 2, p. 1, 2014.
- [18] WeGLARZ, A. Meat quality defined based on ph and colour depending on cattle category and slaughter season. *Czech Journal of Animal Science*, v. 55, n. 12, p. 548–556, 2010.
- [19] CAMPOS, P. F. et al. Impactos da seleção genética na qualidade da carne suína. *PUBVET*, v. 8, n. 2, January 2014.
- [20] USA, P. *Meat Quality: Understanding Industry Measurements and Guidelines*. 2003. <<http://kaslobay.ca/downloads/techupdate/meatquality.pdf>>. Online; acesso em 26 de Dezembro de 2015.
- [21] CHEN, Q. et al. Application of linear/non-linear classification algorithms in discrimination of pork storage time using fourier transform near infrared (ft-nir) spectroscopy. *LWT-Food Science and Technology*, Elsevier, v. 44, n. 10, p. 2053–2058, 2011.
- [22] PRZYBYLAK, A. et al. Marbling classification of lambs carcasses with the artificial neural image analysis. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. *Seventh International Conference on Digital Image Processing (ICDIP15)*. Los Angeles, United States, 2015. p. 963113–963113.
- [23] RAVIKANTH, L. et al. Classification of contaminants from wheat using near-infrared hyperspectral imaging. *Biosystems Engineering*, v. 135, p. 73 – 86, 2015. ISSN 1537-5110.
- [24] SHAN, Z. et al. Study on hyperspectral image technology based on manifold fuzzy clustering for pork quality classification. *Journal of Food Safety and Quality*, Journal of Food Safety and Quality, v. 6, n. 4, p. 1421–1428, 2015.
- [25] ZAPOTOCZNY, P.; SZCZYPIŃSKI, P. M.; DASZKIEWICZ, T. Evaluation of the quality of cold meats by computer-assisted image analysis. *{LWT} - Food Science and Technology*, v. 67, p. 37 – 49, 2016. ISSN 0023-6438.
- [26] QIAO, J. et al. Prediction of drip-loss, ph, and color for pork using a hyperspectral imaging technique. *Meat Science*, v. 76, n. 1, p. 1 – 8, 2007. ISSN 0309-1740.
- [27] LIU, M. et al. Comparison of random forest, support vector machine and back propagation neural network for electronic tongue data classification: Application to the recognition of orange beverage and chinese vinegar. *Sensors and Actuators B: Chemical*, v. 177, p. 970 – 980, 2013. ISSN 0925-4005.

- [28] ARGYRI, A. A. et al. A comparison of raman and ft-ir spectroscopy for the prediction of meat spoilage. *Food Control*, v. 29, n. 2, p. 461 – 470, 2013. ISSN 0956-7135. Predictive Modelling of Food Quality and Safety.
- [29] KAMRUZZAMAN, M. et al. Fast detection and visualization of minced lamb meat adulteration using {NIR} hyperspectral imaging and multivariate image analysis. *Talanta*, v. 103, p. 130 – 136, 2013. ISSN 0039-9140.
- [30] LIU, Y. et al. Nondestructive detection of pork comprehensive quality based on spectroscopy and support vector machine. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. *SPIE Sensing Technology+ Applications*. Baltimore, Maryland, USA, 2014. p. 91080R–91080R.
- [31] ROPODI, A. et al. Multispectral image analysis approach to detect adulteration of beef and pork in raw meats. *Food Research International*, v. 67, p. 12 – 18, 2015. ISSN 0963-9969.
- [32] FACELI, K. et al. Inteligência artificial: Uma abordagem de aprendizado de máquina. *Rio de Janeiro: LTC*, 2011.
- [33] VAIDYA, A. The epistemology of modality. In: ZALTA, E. N. (Ed.). *The Stanford Encyclopedia of Philosophy*. Summer 2014. Stanford, California, USA: Metaphysics Research Lab, Stanford University, 2014. Online; acesso em 11 de Agosto de 2015.
- [34] LIRA, G. A. Estudo da correlação das características clínico-patológicas do câncer colorretal com a expressão imunohistoquímica de proteínas da progressão tumoral. Universidade Federal do Rio Grande do Norte, 2015.
- [35] CHEN, J. C.; LI, C. P. Tone reservation using near-optimal peak reduction tone set selection algorithm for papr reduction in ofdm systems. *IEEE Signal Processing Letters*, v. 17, n. 11, p. 933–936, Nov 2010. ISSN 1070-9908.
- [36] RUSSELL, S. J.; NORVIG, P. *Artificial intelligence: a modern approach*. Upper Saddle River, New Jersey: Pearson, 2009. v. 3.
- [37] RIZA, L. S. et al. frbs: Fuzzy rule-based systems for classification and regression in R. *Journal of Statistical Software*, v. 65, n. 6, p. 1–30, 2015.
- [38] ISHIBUCHI, H.; NAKASHIMA, T.; MURATA, T. Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, v. 29, n. 5, p. 601–618, 1999.
- [39] MARSLAND, S. *Machine Learning: An Algorithmic Perspective*. 1st. ed. New Jersey, USA: Chapman & Hall/CRC, 2009.
- [40] HARRINGTON, P. *Machine learning in action*. Shelter Island, NY, USA: Manning Greenwich, CT, 2012. v. 5.
- [41] QUINLAN, J. R. *C4.5: Programs for Machine Learning*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1993. ISBN 1558602402.
- [42] MITCHELL, T. M. *Machine Learning*. 1. ed. New York, NY, USA: McGraw-Hill, Inc., 1997.

- [43] MAIMON, O.; ROKACH, L. *Data Mining and Knowledge Discovery Handbook*. second. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2010.
- [44] HAYKIN, S. *Neural Networks: A Comprehensive Foundation*. Upper Saddle River, NJ, USA: Prentice Hall, 1999.
- [45] BALAMURUGAN, A. alias et al. Nb+: an improved naive bayesian algorithm. *Knowledge-Based Systems*, Elsevier, v. 24, n. 5, p. 563–569, 2011.
- [46] BREIMAN, L. Random forests. *Machine Learning*, Kluwer Academic Publishers, v. 45, n. 1, p. 5–32, 2001. ISSN 0885-6125.
- [47] BREIMAN, L. Bagging predictors. *Machine Learning*, Kluwer Academic Publishers, v. 24, n. 2, p. 123–140, 1996. ISSN 0885-6125.
- [48] VAPNIK, V. N. *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer-Verlag New York, Inc., 1995. ISBN 0-387-94559-8.
- [49] BEN-HUR, A.; WESTON, J. A user’s guide to support vector machines. In: *Data Mining Techniques for the Life Sciences*. Totowa, NJ: Humana Press, 2010, (Methods in Molecular Biology, v. 609). p. 223–239.
- [50] LAKSHMI, T. J.; PRASAD, C. S. R. A study on classifying imbalanced datasets. In: IEEE. *Networks & Soft Computing (ICNSC), 2014 First International Conference on*. [S.l.], 2014. p. 141–145.
- [51] YU, C. et al. Multi-classes imbalanced dataset classification based on sample information. In: IEEE. *2015 IEEE 17th International Conference on High Performance Computing and Communications, 2015 IEEE 7th International Symposium on Cyberspace Safety and Security, and 2015 IEEE 12th International Conference on Embedded Software and Systems*. [S.l.], 2015. p. 1768–1773.
- [52] KASABOV, N. K. *Foundations of neural networks, fuzzy systems, and knowledge engineering*. [S.l.]: The MIT Press, 1996.
- [53] COPPIN, B. *Inteligência Artificial*. [S.l.]: Grupo Gen-LTC, 2010.
- [54] JENSEN, R.; SHEN, Q. *Computational intelligence and feature selection: rough and fuzzy approaches*. Hoboken, New Jersey: John Wiley & Sons, 2008. v. 8.
- [55] ALCALÁ, R. et al. Techniques for learning and tuning fuzzy rule-based systems for linguistic modeling and their application. *Knowledge Engineering Systems, Techniques and Applications*, Citeseer, v. 3, p. 889–941, 1999.
- [56] IQBAL, A. et al. Classification of pre-sliced pork and turkey ham qualities based on image colour and textural features and their relationships with consumer responses. *Meat Science*, v. 84, n. 3, p. 455 – 465, 2010. ISSN 0309-1740.
- [57] VALOUS, N. A. et al. Supervised neural network classification of pre-sliced cooked pork ham images using quaternionic singular values. *Meat Science*, v. 84, n. 3, p. 422 – 430, 2010. ISSN 0309-1740.
- [58] MONROY, M. et al. Pork meat quality classification using visible/near-infrared spectroscopic data. *Biosystems Engineering*, v. 107, n. 3, p. 271 – 276, 2010. ISSN 1537-5110.

- [59] LIU, L. et al. Categorization of pork quality using gabor filter-based hyperspectral imaging technology. *Journal of Food Engineering*, v. 99, n. 3, p. 284 – 293, 2010. ISSN 0260-8774.
- [60] BARBIN, D. et al. Near-infrared hyperspectral imaging for grading and classification of pork. *Meat Science*, v. 90, n. 1, p. 259 – 268, 2012. ISSN 0309-1740.
- [61] MA, J. et al. Application of vis–nir hyperspectral imaging in classification between fresh and frozen-thawed pork longissimus dorsi muscles. *International Journal of Refrigeration*, v. 50, p. 10 – 18, 2015. ISSN 0140-7007.
- [62] BARBON, A. P. A. et al. Storage time prediction of pork by computational intelligence. *Computers and Electronics in Agriculture*, v. 127, p. 368 – 375, 2016. ISSN 0168-1699.
- [63] TAO, F.; PENG, Y. A method for nondestructive prediction of pork meat quality and safety attributes by hyperspectral imaging technique. *Journal of Food Engineering*, v. 126, p. 98 – 106, 2014. ISSN 0260-8774.
- [64] BALAGE, J. M. et al. Predicting pork quality using vis/nir spectroscopy. *Meat Science*, v. 108, p. 37 – 43, 2015. ISSN 0309-1740.
- [65] PRIETO, N. et al. Rapid discrimination of enhanced quality pork by visible and near infrared spectroscopy. *Meat Science*, v. 110, p. 76 – 84, 2015. ISSN 0309-1740.
- [66] DISSING, B. S. et al. Using multispectral imaging for spoilage detection of pork meat. *Food and Bioprocess Technology*, Springer, v. 6, n. 9, p. 2268–2279, 2013.
- [67] PU, H. et al. Classification of fresh and frozen-thawed pork muscles using visible and near infrared hyperspectral imaging and textural analysis. *Meat Science*, v. 99, p. 81 – 88, 2015. ISSN 0309-1740.
- [68] BAUER, A. et al. Towards an online assessment of meat quality in pork. *59th International Congress of Meat Science and Technology*, August 2013.
- [69] ADORNI, G.; BIANCHI, D.; CAGNONI, S. Learning fuzzy decision trees for ham quality control. *Trans. QCAV2001*, p. 259–264, 2001.
- [70] WARRISS, P. D.; BROWN, S. N. The relationship between initial ph, reflectance and exudation in pig muscle. *Meat Science*, v. 20, p. 65–74, 1987.
- [71] LAACK, R. L. J. V. et al. Is colour brightness (l-value) a reliable indicator of water-holding capacity in porcine muscle? *Meat Science*, v. 38, p. 193–201, 1994.
- [72] FAUCITANO, L. et al. Shelf life of pork from five different quality classes. *Meat Science*, v. 84, n. 3, p. 466 – 469, 2010. ISSN 0309-1740.
- [73] KAUFFMAN, R. et al. The effectiveness of examining early post-mortem musculature to predict ultimate pork quality. *Meat Science*, Elsevier, v. 34, n. 3, p. 283–300, 1993.
- [74] JOO, S. et al. Objectively predicting ultimate quality of post-rigor pork musculature. *Asian-Australasian Journal of Animal Sciences*, v. 13, n. 1, p. 77–85, 2000.

- [75] BARBUT, S. Estimates and detection of the pse problem in young turkey breast meat. *Canadian Journal of Animal Science*, v. 76, 1996.
- [76] BOCCARD R.; BUCHTER, L. C. E. e. a. Procedures for measuring meat quality characteristics in beef production experiments. report of a working group in the commission of the european communities' (cec) beef production research programme. *Livestock Production Science*, v. 8, 1981.
- [77] PERES, L. M.; OLIVEIRA, E. R. de; AL, C. L. de Lucio; et. Análise comparativa de metodologias de determinação da capacidade de retenção de água. In: *XXI Congresso Brasileiro de Zootecnia*. Maceió, Brasil: [s.n.], 2011.
- [78] FAUCITANO, L. Effects of preslaughter handling on the pig welfare and its influence on meat quality. In: *Conferência Virtual Internacional sobre Qualidade de carne Suína*. Concórdia, SC, Brasil: [s.n.], 2000. v. 16.
- [79] CHEVILLON, P. O bem-estar dos suínos durante o pré-abate e no atordoamento. In: *i conferência virtual internacional sobre qualidade de carne suína*. [S.l.: s.n.], 2000. v. 16.
- [80] MAGANHINI, M. B. et al. Carnes pse (pale, soft, exudative) e dfd (dark, firm, dry) em lombo suíno numa linha de abate industrial. *Ciência e Tecnologia de Alimentos*, Sociedade Brasileira de Ciência e Tecnologia de Alimentos, v. 27, n. 1, p. 69–72, 2007.
- [81] RIZA, L. S. et al. Learning from data using the r package “frbs”. In: *2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. Beijing, China: [s.n.], 2014. p. 2149–2155.
- [82] FRIEDMAN, M. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the american statistical association*, Taylor & Francis, v. 32, n. 200, p. 675–701, 1937.
- [83] NEMENYI, P. Distribution-free multiple comparisons. In: INTERNATIONAL BIOMETRIC SOC 1441 I ST, NW, SUITE 700, WASHINGTON, DC 20005-2210. *Biometrics*. [S.l.], 1962. v. 18, n. 2, p. 263.

TRABALHOS PUBLICADOS PELO AUTOR

Trabalhos publicados pelo autor durante o programa.

1. Barbon, A.P.A.C; Mantovani, R.G.; Fuzyi, E.M.; Peres, L.M.; Bridi, A.M.; Barbon Jr, S., **Storage time prediction of pork by Computational Intelligence**, Computers and Electronics in Agriculture, 2016, Elsevier, 368 - 375, (Qualis CC 2012, A2)
2. Fuzyi, E.M.; da Silva, M.C. ; Barbon JR., S., **Solução M-Health no pré-diagnóstico da Laringe**, Simpósio Brasileiro de Sistemas de Informação (SBSI), 2015, ACM, 67:501–67:508, (Qualis CC 2012, B4)
3. Fuzyi, E.M.; Armani, A. ; Barbon JR., S., **Modelagem de banco de dados para diagnóstico de transtornos mentais pela voz**, Journal of Health Informatics, 2015, 23-29, (Qualis CC 2012, B5)