

unesp



UNIVERSIDADE ESTADUAL PAULISTA

Programa de Pós-Graduação em Engenharia Elétrica

**CARACTERIZAÇÃO DE LEITE
UTILIZANDO TÉCNICAS DE ULTRA-SOM
E REDES NEURAIIS**

SÉRGIO LUIZ SOUSA NAZARIO

MESTRADO

Ilha Solteira - SP



UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Campus de Ilha Solteira

Programa de Pós-Graduação em Engenharia Elétrica

Caracterização de leite utilizando técnicas de ultra-som e redes neurais

Sérgio Luiz Sousa Nazario

Orientador: Prof. Dr. Ricardo Tokio Higuti

Dissertação apresentada à Faculdade de Engenharia - UNESP - Campus de Ilha Solteira para obtenção do título de Mestre em Engenharia Elétrica.

Área de Concentração: Automação

Ilha Solteira - SP

novembro/2007

FICHA CATALOGRÁFICA

Elaborada pela Seção Técnica de Aquisição e Tratamento da Informação
Serviço Técnico de Biblioteca e Documentação da UNESP - Ilha Solteira.

Nazario, Sérgio Luiz Sousa
N335c Caracterização de leite utilizando técnicas de ultra-som e redes neurais / Sérgio Luiz
Sousa Nazario. -- Ilha Solteira : [s.n.], 2007
101p. : il.

Dissertação (mestrado) - Universidade Estadual Paulista. Faculdade de Engenharia de
Ilha Solteira. Área de conhecimento: Automação, 2007

Orientador: Ricardo Tokio Higuti
Bibliografia: p. 98-101

1. Ultra-som. 2. Leite. 3. Redes neurais (Computação) 4. Caracterização de líquidos.

Agradecimentos

A Deus, pela chance que nos dá de aprender, desenvolver e passar o conhecimento a cada dia.

Ao Prof. Dr. Ricardo Tokio Higuti, pelo empenho, dedicação, confiança na realização do meu trabalho e pela boa vontade em transferir seus conhecimentos.

À Prof. Dr. Jacira dos Santos Isepon, pelo apoio, dedicação, e pela boa vontade em transferir seus conhecimentos.

Aos meus pais, Ailton e Veneranda, que sempre me incentivaram e apoiaram na realização de um sonho. Aos meus irmãos, Ailton Junior e Carlos Augusto, pelo carinho e apoio dedicado.

À minha madrinha Zoraide e meu padrinho Delionizio por toda a ajuda.

À minha namorada Hérica, pelo amor, apoio, compreensão, incentivo e ajuda em todos os momentos.

Aos meus colegas que estiveram ao meu lado em cada caminhada, em especial Luiz Antonio Marçal, André Luiz, Marcelo Sanches, Wander Wagner, João Marcos Sakamoto e Wesley Pontes.

Aos amigos do Laboratório de Ultra-som, Guilherme, Rafael Batista, Vander, Ricardo, Rafael Ozaki pelas trocas de idéias e momentos de descontração e alegria.

Aos professores Cláudio Kitano e Aparecido Augusto de Carvalho pelo apoio, dedicação e amizade e a todos os professores do DEE que direta ou indiretamente contribuíram para este trabalho.

Aos técnicos do laboratório de Biotecnologia de alimentos, Valdeci e Antonia, por sua ajuda e dedicação.

A todos meus amigos de pós graduação pelas horas de estudo, dedicação e amizade.

Agradeço acima de tudo a Deus, que me concedeu força e perseverança para continuar caminhando sempre em busca de meus sonhos.

*“Sábio é o ser humano que tem coragem de ir
diante do espelho da sua alma para reconhecer
seus erros e fracassos e utilizá-los para plantar
as mais belas sementes no terreno de sua in-
teligência.”*

Augusto Cury

NAZARIO, Sérgio Luiz Sousa. **Caracterização de leite utilizando técnicas de ultra-som e redes neurais**. 2007. 101 f. Dissertação (Mestrado em Engenharia Elétrica - Automação) - Faculdade de Engenharia, Universidade Estadual Paulista, Ilha Solteira, 2007.

Resumo

Na indústria de laticínios, a qualidade do leite é medida por alguns parâmetros como gordura, lactose, proteína e água adicionada. Baixos teores de gordura podem diminuir suas propriedades nutricionais, e o que pode ser um indicativo de adulteração, com conseqüências econômicas e para a saúde. Por estas razões, o leite recebido de produtores deve passar por alguns testes antes de sua aceitação nos laticínios. Atualmente, as técnicas utilizadas para medição de parâmetros do leite em geral envolvem manipulação de produtos químicos e podem ser demoradas e destrutivas. Desta forma, observa-se a necessidade do desenvolvimento de métodos rápidos, de fácil manipulação e que não tragam perigo ao ser humano. Este trabalho apresenta um estudo sobre a caracterização de leite fluido utilizando técnicas de ultra-som e redes neurais artificiais. Métodos de ultra-som podem ser rápidos, não-destrutivos e não-agressivos enquanto que redes neurais possuem a vantagem de aprender através de exemplos apresentados e de generalizar as informações aprendidas. As redes neurais também são capazes de extrair informações não apresentadas de forma explícita, através de exemplos. Utiliza-se uma célula de medição de parâmetros acústicos para obter dados como a velocidade de propagação, coeficiente de atenuação e densidade de amostras. Esses valores são correlacionados com as propriedades do leite, como teor de gordura e teor de água adicionada, que são medidos com métodos convencionalmente utilizados em laticínios e laboratórios de tecnologia de alimentos, como o método Gerber e o crioscópio. Esses dados são utilizados para projetar uma rede neural artificial do tipo MLP (*Multi-Layer Perceptron*), que fornece na saída o teor de gordura e a quantidade de água adicionada ao leite, a partir dos parâmetros medidos pela célula de medição. As redes neurais desenvolvidas resultaram em mais de 95 % de amostras classificadas corretamente, com uma resolução de 0,1 % na determinação da quantidade de gordura. Para quantidade de água adicionada, a resolução foi de 1 % para 1 a 10 % de água adicionada e 5 % para quantidade de água adicionada de 10 a 60 %.

Palavras-chave: ultra-som, leite, redes neurais, caracterização de líquidos.

NAZARIO, Sérgio Luiz Sousa. **Characterization of milk using ultrasonic and neural networks techniques**. 2007. 101 p. Dissertation (Master in Electrical Engineering - Automation) - Faculdade de Engenharia, Universidade Estadual Paulista, Ilha Solteira, 2007.

Abstract

In the dairy industry, quality of milk is measured by some parameters such as fat, lactose, protein and water content (or added water). Lower values of fat content can decrease its nutritional properties, while higher values of water content may be an indicative of adulteration, for example, resulting in health and economic impacts. For these reasons, in dairy industries, milk that is received from producers must be tested or certificated before its acceptance. Nowadays, the techniques used for milk parameters measurement, in general, involve manipulation of chemical products, may be time-consuming and destructive. Consequently, the development of faster methods, with ease of manipulation and that are not harmful to the human being are important. This work presents a study about the characterization of milk using ultrasound techniques and neural networks. Ultrasound methods can be fast, non destructive and non aggressive, while neural networks techniques have the advantage of learning through presented examples and of generalizing the learned information. An ultrasonic liquid measurement cell is used to obtain data as acoustic propagation velocity, attenuation coefficient and density of samples. Those values are correlated with milk properties, as fat content and added water, that are measured with conventional methods used in dairy industries and food technology laboratories, such as the Gerber method and the crioscope. Using these data, MLP (Multi-Layer Perceptron) neural networks are obtained, that determine the fat content and the amount of added water to the milk, having as inputs the measured acoustic parameters from the measurement cell. The developed neural networks resulted in more than 95% of correct classifications, with a resolution of 0,1% in the determination of the fat content. For added water, the resolution was 1% for 1 to 10% of added water and 5% in the range of 10 to 60% of added water.

Keywords: ultrasonics, milk, neural networks, liquid characterization.

Lista de Figuras

2.1	Esquema para medição da densidade: Proveta e Termolactodensímetro.	25
2.2	Centrífuga de Gerber e butirômetros.	27
2.3	Crioscópio.	28
3.1	Incidência normal de uma onda a uma interface.	35
3.2	Célula de medição.	36
3.3	Foto da célula de medição: vista superior.	37
3.4	Foto da célula de medição: TDE à esquerda e à direita, refletor e câmara de amostra.	37
3.5	Gráfico dos sinais de eco no domínio do tempo.	39
3.6	Gráfico dos sinais recebidos no domínio da frequência.	39
4.1	Neurônio artificial de McCulloch e Pitts.	45
4.2	Função de ativação linear.	45
4.3	Função de ativação rampa.	46
4.4	Função degrau.	46
4.5	Função de ativação sigmoidal.	47
4.6	Aprendizado supervisionado.	48
4.7	Aprendizado não-supervisionado.	48
4.8	Neurônio artificial.	51
4.9	Arquitetura perceptron de múltiplas camadas com uma camada oculta.	51
4.10	Ilustração das direções de dois fluxos básicos em um perceptron de múltiplas camadas: Propagação para frente e retropropagação.	52
4.11	Ilustração da regra de parada antecipada baseada na validação cruzada.	62
4.12	<i>a</i> -Dados ajustados adequadamente; <i>b</i> -Dados ajustados em excesso.	63
5.1	Concentração de gordura em função da quantidade de água adicionada.	70
5.2	Esquema experimental.	71
5.3	Foto do sistema de medição	72

5.4	Velocidade de propagação da onda no leite UHT em função da temperatura para diferentes concentrações de gordura.	74
5.5	Velocidade de propagação da onda no leite UHT, ampliada, em função da temperatura para diferentes concentrações de gordura.	74
5.6	Velocidade de propagação da onda no leite UHT em função da concentração de gordura.	75
5.7	Coefficiente de atenuação da onda no leite UHT para frequência de 5 MHz em função do teor de gordura.	76
5.8	Coefficiente de atenuação da onda no leite UHT em função da frequência, para diferentes concentrações de gordura.	76
5.9	Coefficiente de atenuação da onda no leite UHT para frequência de 5 MHz em função da temperatura.	77
5.10	Densidade do leite UHT em função da concentração de gordura. . .	77
5.11	Densidade do leite UHT com 3,0% de teor de gordura (ϕ_v) em função da Temperatura.	78
5.12	Velocidade de propagação da onda no leite UHT com diferentes quantidades de água adicionada em função da temperatura.	78
5.13	Velocidade de propagação da onda no leite UHT em função de água (H_2O) adicionada.	79
5.14	Coefficiente de atenuação da onda no leite UHT em função da quantidade de água adicionada	80
5.15	Coefficiente de atenuação da onda no leite UHT para diversas frequências em função da quantidade de água adicionada	80
5.16	Densidade para o leite UHT integral em função de água adicionada.	81
5.17	Densidade do leite UHT integral, para os métodos convencionais (Termolactodensímetro e picnômetro) e por ultra-som ($f=5MHz$) em função de água adicionada.	82
5.18	Esquema da rede neural projetada para determinação do teor de gordura, com cinco entradas (velocidade de propagação, temperatura, coeficiente de atenuação em três frequências) e 26 saídas, com resolução igual a 0,1%.	82
5.19	Esquema da rede neural projetada para determinação do teor de água, com cinco entradas (velocidade de propagação, temperatura, coeficiente de atenuação em três frequências) e 21 saídas, com resolução igual a 1% (entre 1 e 10%) e 5% (entre 10 e 60%).	83
5.20	Capacidade de classificação usando três e cinco nodos na camada de entrada para detecção do teor de gordura.	89
5.21	Capacidade de classificação usando cinco nodos na camada de entrada para detecção da quantidade de água adicionada ao leite.	90

5.22	Decaimento do erro de treinamento (linha sólida) e validação (linha tracejada), para detecção do teor de gordura.	92
------	---	----

Lista de Tabelas

2.1	Componentes principais do leite em média para o leite bovino. . . .	22
2.2	Ponto crioscópico versus % de água.	29
5.1	Informação nutricional: Porção de 200 ml para o leite UHT integral.	68
5.2	Teor de gordura (ϕ_v) obtido pelo método de Gerber, em temperatura ambiente (26°C) para diluições de leite UHT em porcentagens diferentes.	69
5.3	Teor de gordura (ϕ_v) obtido pelo método de Gerber, em temperatura ambiente (26°C) para o leite UHT integral com diferentes porcentagens de água adicionada.	69
5.4	Densidades obtidas pelos métodos que utilizam o termolactodensímetro e picnômetro para diferentes concentrações de gordura à temperatura média de 26°C.	70
5.5	Densidades obtidas pelos métodos que utilizam o termolactodensímetro e picnômetro, para diferentes quantidades de água à temperatura média de 25°C	71
5.6	Densidades obtidas pelos métodos de ultra-som, Termolactodensímetro e picnômetro para variação de teor de água adicionada (a.a.) . . .	81
5.7	Redes treinadas para a topologia com três camadas de entrada para detecção do teor de gordura.	87
5.8	Redes treinadas para a topologia com cinco camadas de entrada para detecção do teor de gordura.	88
5.9	Redes treinadas para a topologia com cinco camadas de entrada para detecção de água adicionada.	90

Lista de Símbolos

Símbolos

$a_i(t)$	sinal de eco
$A_i(f)$	espectro do sinal de eco $a_i(t)$
α_i	coeficiente de atenuação no meio i [Np/m = 8,686 dB/m]
c_i	velocidade de propagação no meio i [m/s]
f	frequência [Hz]
k	constante de propagação ($= \omega / c + i \alpha$)
L_2	tamanho da câmara de amostra [m]
R_{ij}	coeficiente de reflexão entre os meio i e j
T_{ij}	coeficiente de transmissão entre os meio i e j
Z	impedância acústica [Rayls]
β	coeficiente de expansão volumétrico [$1/^\circ C$]
Δt	diferença de tempo [s]
ϕ_v	concentração em volume
t.g.	teor de gordura
a.a.	quantidade de água adicionada
ρ	densidade [kg/m ³]
ω	frequência angular [rad/s]
$e_\kappa(n)$	erro na camada de saída do neurônio
$\epsilon(n)$	soma dos erros quadráticos
$w_{ji}(n)$	peso sináptico
$d_\kappa(n)$	resposta desejada
$y_j(n)$	função de ativação que aparece na saída do neurônio
δ_κ	gradiente local
SSE	soma de quadrados dos erros
α	constante de momento
η	taxa de aprendizado

Sumário

1	Introdução e Justificativas	16
1.1	Objetivos	19
1.2	Organização do trabalho	19
2	O leite e técnicas de análise convencionais	21
2.1	Principais componentes do leite	21
2.2	Obtenção higiênica do leite	22
2.3	O leite como líquido não-homogêneo	24
2.4	Técnicas de análise de leite fluido	25
2.4.1	Determinação da densidade	25
2.4.2	Determinação de teor de gordura	26
2.4.3	Determinação do ponto de congelamento do leite	28
2.5	Comentários	29
3	Técnicas de ultra-som e célula de medição	31
3.1	Parâmetros acústicos	31
3.1.1	Velocidade de propagação	31
3.1.2	Coeficiente de atenuação	32
3.1.3	Impedância acústica e densidade	34
3.1.4	Fenômenos de transmissão	34
3.2	Célula de medição	36
3.2.1	Coeficiente de reflexão	38
3.2.2	Velocidade de propagação	40
3.2.3	Densidade e coeficiente de atenuação	41
3.3	Comentários	42
4	Redes neurais artificiais	43
4.1	Redes neurais artificiais (RNAs)	43
4.2	Neurônio biológico	44
4.3	Neurônio artificial de McCulloch e Pitts	44
4.4	Funções de ativação	45
4.5	Aprendizagem	47

4.5.1	Aprendizado supervisionado	47
4.5.2	Aprendizado não-supervisionado	48
4.5.3	Aprendizado por competição	48
4.5.4	Aprendizado por reforço	48
4.6	Redes perceptron e adaline	49
4.7	Redes MLP (<i>Multi Layer Perceptron</i>)	49
4.7.1	Arquitetura do perceptron de múltiplas camadas	50
4.7.2	Treinamento de redes multicamadas	51
4.7.3	Algoritmo de retropropagação	51
4.8	Modo de treinamento seqüencial e por lote	55
4.8.1	Vetor gradiente	56
4.8.2	Gradiente descendente	56
4.8.3	Modo por lote	58
4.9	Taxa de aprendizado e constante de momento	58
4.10	Função de ativação	59
4.11	Crítério de parada	59
4.12	Representação da saída e regra de decisão	60
4.13	Divisão do conjunto de dados	61
4.14	Validação cruzada	61
4.15	Generalização	62
4.15.1	Avaliação da eficiência de um algoritmo de aprendizagem	63
4.16	Heurística para melhorar o desempenho	64
4.17	Técnica de poda de rede	65
4.18	Virtudes e limitações	65
4.18.1	Vantagens	65
4.18.2	Desvantagens	65
4.19	Comentários	66
5	Métodos experimentais e resultados	67
5.1	Preparação e calibração de amostras	67
5.2	Descrição do equipamento de ultra-som	70
5.3	Calibração da célula	72
5.4	Resultados: métodos de ultra-som	73
5.4.1	Variação do teor de gordura	73
5.4.2	Variação da quantidade de água adicionada	77
5.5	Comparação com métodos convencionais	79
5.6	Métodos e resultados: redes neurais	81
5.6.1	Planejamento dos experimentos	84
5.6.2	Procedimento e descrição dos parâmetros	85
5.6.3	Execução dos experimentos	86
5.6.4	Resultados	86

5.6.5	Dificuldades encontradas e limitações do modelo	92
5.7	Comentários	92
6	Conclusões	95
6	Referências Bibliográficas	98

Capítulo 1

Introdução e Justificativas

Um setor importante no cenário agroindustrial brasileiro é a indústria leiteira. Apesar das grandes dificuldades dos produtores em adquirir recursos para melhoramento genético de seus rebanhos, melhorar sua estrutura de produção e ter acesso a novas tecnologias, este setor vem tendo um enorme crescimento ao longo do tempo. Nos últimos anos, houve um crescimento anual de cerca de 3% em produção, superando assim o valor bruto da produção agroindustrial de setores de grande importância, como arroz e café beneficiado [1]. Relatos indicam que uma grande área da pastagem brasileira ainda não foi utilizada pelo setor, ao passo que as demais pastagens se encontram degradadas e não são utilizadas de forma eficiente. Através de investimentos no setor, podem ser alcançados melhores rendimentos e, conseqüentemente, obter resultados que contribuirão para o desenvolvimento sócio-econômico brasileiro [2].

O setor leiteiro também é responsável por uma arrecadação de 2,1 bilhões de reais, através de impostos, de forma direta e indireta. O crescimento deste setor também contribui para a geração de empregos no campo, na indústria de laticínios, e um aumento no consumo de insumos e de equipamentos. O Instituto Brasileiro de Geografia e Estatística (IBGE) contabilizou cerca de 3338 laticínios no Brasil em 2001, que pagaram cerca de R\$ 768 milhões em salários e remunerações, superando setores de grande importância como o de eletrodomésticos e têxtil [1].

Apesar de ser um dos maiores produtores de leite mundial, o Brasil enfrenta uma grande concorrência com produtos oriundos de países da América do Sul devido à consolidação do MERCOSUL e da qualidade dos produtos, tornando o Brasil um grande importador de produtos derivados do leite [3]. Em paralelo a estas informações, outro dado que reforça a necessidade para um desenvolvimento tecnológico e de pessoal capacitado, foi um aumento considerável no consumo de leite fluido, chegando a uma cifra de 68,4 kg/pessoa/ano em 2004 [3].

O leite é um dos alimentos mais importantes devido à sua composição nutricional e exerce um papel muito importante para o desenvolvimento e formação das crianças, além de ter um papel importantíssimo para redução de doenças na popu-

lação em geral, como osteoporose, problemas na formação óssea e sua manutenção. Atualmente existe um programa social do governo que tem como objetivo o aumento na produção por pequenos produtores rurais locais e no consumo de leite pela população, gerando crescimento do setor e contribuindo para a redução dos índices de fome no país.

As normas para qualidade do leite foram alteradas pelo Ministério da Agricultura, Pecuária e Abastecimento, em setembro de 2002, através da normativa 51 [4]. Estas novas normas têm como objetivo melhorar a qualidade dos produtos oferecidos aos consumidores no Brasil e no exterior. O conjunto de normas visam cuidados fundamentais para a boa qualidade do leite fluido, presentes em todas as etapas, desde a produção, armazenagem, o transporte, a industrialização e a comercialização. Hoje, na maioria dos laticínios, as análises da qualidade do leite são realizadas por processos químicos ou convencionais, que são demorados, destrutivos (amostra inaproveitável após a medição) e caros. Além disso, são utilizados produtos que oferecem perigo à saúde humana. Então, é clara a necessidade de desenvolvimento de novas técnicas que possam analisar o leite com maior rapidez, precisão e menor custo.

O investimento em tecnologias e pessoal qualificado, para desenvolvimento de novos produtos derivados do leite, são de grande importância para o crescimento da produção, pois agregam valor ao produto final. O desenvolvimento de novos equipamentos que possam monitorar com maior facilidade e maior exatidão a matéria-prima vinda do produtor rural são aspectos importantes que podem contribuir para controlar os padrões exigidos por órgãos governamentais em relação à qualidade nutricional do produto e para inibir o processo de adulteração, por parte de alguns produtores mal intencionados que praticam o ato criminoso de falsificação do leite por adição de água, leite desnatado e soro do leite [5]. O objetivo da adulteração é de aumentar a quantidade de leite entregue ao laticínios o que proporciona ganho de produção.

Uma alternativa a ser considerada é o uso de ultra-som para caracterizar líquidos. Aplicações de ultra-som na área de alimentos podem ser distinguidas a partir de técnicas de alta e baixa potência [6]. As técnicas de alta potência são em geral de baixa frequência (20 a 100 kHz), podendo implicar em modificações físico-químicas do material. As técnicas de alta frequência (acima de 500 kHz) e baixa potência são utilizadas para caracterização de alimentos através da determinação de concentração, densidade, tamanho de partícula, fluxo, dimensões entre outras aplicações [7, 8, 9]. Algumas aplicações especiais do ultra-som no setor de laticínios são encontradas na literatura para determinar o teor de gordura [10, 11] e existem estudos de métodos para determinar a coagulação de leite para fabricação de queijo [12]. Não foram encontradas, na literatura, aplicações envolvendo a medição de proteínas, lactose, quantidade de água adicionada, falsificação por adição de soro do leite por técnicas

de ultra-som.

Pela medição da velocidade de propagação acústica e coeficiente de atenuação, pode-se obter indiretamente uma característica como a concentração em emulsões [13, 14]. Métodos de ultra-som são relativamente simples, não-destrutivos e podem ser usados *on-line*.

Miles, Shore e Langley [15] investigaram o fenômeno da atenuação acústica no leite proveniente de dois efeitos: perdas intrínsecas e perdas ligadas às partículas de gordura presentes no meio. Concluíram que as perdas causadas por essas partículas são proporcionais à concentração de gordura presente na amostra.

Bacaneli [10] e Higuti [13] verificaram a relação entre parâmetros acústicos como velocidade de propagação da onda, atenuação e densidade, em função do teor de gordura presente no leite. Foram obtidas curvas de calibração, que foram utilizadas para relacionar esses parâmetros acústicos com as características do leite em função da temperatura. Uma limitação em se trabalhar somente com curvas de calibração pode surgir devido à possibilidade de uma amostra de leite adulterado e uma amostra de leite sem adulteração possuírem um mesmo valor de velocidade de propagação devido a diferentes teores de gordura, por exemplo, o que pode provocar caracterizações erradas. Dessa forma, o uso de técnicas mais elaboradas para a análise dos resultados é necessária.

Dessa forma, uma dificuldade é relacionar os parâmetros acústicos com uma propriedade particular do material analisado. No caso de leite, o conteúdo de gordura pode ser alterado através de vários fatores: tipo de leite e criação do gado, conteúdo de água e adulteração, por exemplo. Por conseguinte há uma necessidade em desenvolver e aplicar novas técnicas de análise para a classificação de amostras. Entre elas, redes neurais têm sido aplicadas em alguns problemas.

Redes Neurais Artificiais (RNAs) são sistemas paralelos distribuídos compostos por unidades de processamento simples, chamados neurônios, que calculam determinadas funções matemáticas. Na maioria dos modelos estas conexões estão associadas a pesos, os quais armazenam o conhecimento representado no modelo e servem para ponderar a entrada recebida por cada neurônio da rede [16]. Essas ferramentas robustas e poderosas para o reconhecimento de padrões ou classificação de dados vêm sendo utilizadas com grande poder por pesquisadores, com o intuito de dar suporte a decisões médicas, como por exemplo tratamento de câncer e doenças do coração, análise de imagens de ultra-som, radiografia e mamografia. Yan [17] e Ergun [18] aplicaram técnicas que utilizam redes neurais para diagnosticar doenças no coração de humanos. Tian e Shang [19] utilizaram redes neurais para classificar espécies diferentes de ratos. Ochoa [20] utilizou redes neurais para classificar e reconhecer sistemas de navegação e mapeamento de rotas. Já Jian *et al* [21] usaram redes neurais para caracterizar e avaliar nanoestruturas.

Redes neurais artificiais são uma excelente alternativa para resolução de prob-

lemas de classificação, uma vez que o processamento é estruturalmente paralelo, garantindo um desempenho superior ao dos modelos convencionais, possibilitando diversas funcionalidades como: adaptabilidade, tolerância à falha e abstração, aliando-as à velocidade de resposta, entre outros. Além disso, o que torna as redes neurais uma técnica interessante para o problema é a capacidade de ser projetada para generalizar corretamente e produzir um mapeamento de entrada-saída correto, mesmo quando a entrada for um pouco diferente dos exemplos usados para treinar a rede [16].

Diante do contexto, é proposta uma investigação sobre a relação entre os parâmetros acústicos e características do leite, utilizando técnicas de medição de propriedades acústicas em líquidos e técnicas de reconhecimento de padrões empregando redes neurais.

1.1 Objetivos

O trabalho tem como objetivo a aplicação de técnicas de ultra-som e redes neurais para caracterização de leite fluido, em termos de gordura e adulteração por adição de água. Serão determinados limites mínimos e resolução, na medição desses parâmetros por ultra-som, para efeito de comparação e validação com métodos convencionais utilizados em laticínios.

1.2 Organização do trabalho

No capítulo 2, são apresentadas as características dos componentes principais que formam o leite, o processo de obtenção do leite no meio rural, a forma com que ele é tratado e subdividido de acordo com seu tratamento e as técnicas convencionais de medição dos parâmetros do leite disponíveis nos laboratórios da Faculdade de Engenharia de Ilha Solteira (FEIS-UNESP). Essas técnicas são utilizadas para calibrar as amostras, para posterior caracterização por ultra-som.

O capítulo 3 apresenta os parâmetros acústicos principais utilizados neste trabalho, de forma breve mostra as teorias básicas para a caracterização de leite por ultra-som e as técnicas de ultra-som e métodos de análise utilizadas para caracterizar o leite. É apresentada a célula de medição usada no trabalho e descrevem-se os parâmetros que são medidos.

O capítulo 4 apresenta conceitos sobre redes neurais artificiais, técnicas de aprendizado e algumas redes específicas. São abordadas as redes neurais MLP, utilizadas neste trabalho, e as técnicas de projeto das mesmas.

No capítulo 5, são apresentados os métodos experimentais e os resultados dos experimentos realizados por técnicas convencionais, por ultra-som e redes neurais.

Os resultados são relacionados e curvas características são levantadas. Por fim as conclusões finais são apresentadas no capítulo 6.

Capítulo 2

O leite e técnicas de análise convencionais

Neste capítulo são abordadas algumas das principais características do leite e as técnicas convencionais disponíveis no Laboratório de Biotecnologia da Faculdade de Engenharia-UNESP-Campus de Ilha Solteira, utilizadas para caracterização de seus componentes. Os resultados são utilizados para se obter amostras devidamente calibradas para comparação com os parâmetros medidos por ultra-som.

2.1 Principais componentes do leite

O leite pode ser descrito sob o ponto de vista biológico e físico-químico como um líquido constituído por água, gordura, proteínas, lactose e minerais. Pode também ser definido como uma reunião de algumas substâncias na água [22]:

1. suspensão coloidal de pequenas partículas (micelas de caseína ligadas a cálcio e fósforo);
2. emulsão de glóbulos de gordura do leite e vitaminas lipossolúveis, que se encontram em suspensão;
3. solução de lactose, proteínas solúveis em água, sais minerais e vitaminas.

De acordo com os órgãos vigentes de fiscalização do governo, o leite deve ser um produto fresco, normal, integral oriundo de ordenha completa e ininterrupta de vacas saudáveis alimentadas e tratadas de forma conveniente, entre 30 dias antes e 10 dias após o parto [4].

A constituição do leite pode variar de animal para animal e de raça para raça, tendo como um outro fator de grande importância as variações climáticas e alimentação [5, 22].

A qualidade do leite está fortemente ligada às condições higiênicas, pois depende da forma como a ordenha é realizada, da temperatura a qual é mantido após a ordenha e da saúde dos animais.

A composição média do leite, de diferentes raças, segundo Fleishmann [5], é definida com maiores detalhes através da Tabela 2.1. A matéria seca total é a diferença entre toda a matéria do leite menos a água.

Tabela 2.1. Componentes principais do leite em média para o leite bovino.

Componente	Porcentagem
Água	87,5%
Matéria gorda (lipídios)	3,6%
Caseína (proteínas)	3,0%
Albumina (proteínas)	0,6%
Lactose (açúcar)	4,6%
Sais minerais (cinzas)	0,7%

O componente de maior porcentagem no leite, a água, é de grande importância para sua análise, pois influi suavemente em sua densidade. Por ser o componente de maior presença no leite, a determinação da quantidade de água assume um papel relevante no processo de análise. A concentração, a densidade e o teor de gordura sofrem alterações devido à variação da quantidade de água no leite. Desta forma é possível detectar fraudes, por adição de água e soro do leite ao leite *in natura*.

O teor de gordura do leite é um dos elementos de maior importância na composição do leite. A gordura presente no leite, de menor densidade que a água, flutua livremente quando o leite está em repouso. Ele é constituído por partículas de gordura de inúmeros tamanhos que se encontram em suspensão no líquido. A visualização da quantidade e da forma que estas partículas estão aglomeradas no líquido é de grande importância para o processo de caracterização por ultra-som, porque podem influenciar na obtenção de alguns parâmetros acústicos. A lactose é outro importante componente do leite que participa com uma porcentagem de 4,6%.

2.2 Obtenção higiênica do leite

A obtenção do leite de forma higiênica é um processo de grande importância para a valorização do produto. O mercado consumidor, a cada dia, se torna mais exigente e cobra produtos de melhor qualidade. A obtenção do leite é a etapa de maior vulnerabilidade para ocorrência de contaminação do leite, seja por sujeiras presentes no local da ordenha, equipamentos utilizados, microorganismos e substâncias químicas. Essas situações causam a desvalorização do produto ou até mesmo o

não aproveitamento da produção, que pode se converter em prejuízos ao produtor [4].

A Instrução Normativa 51, de 18 de setembro de 2002, veio com o objetivo de aperfeiçoar e modernizar a legislação sanitária federal sobre a produção de leite. As normas entraram em vigor em julho de 2005, nos estados do sul, sudeste e centro-oeste. Para os demais estados, a partir de julho de 2007. Essa instrução normativa estabelece índices mais rígidos de contagem de células somáticas, contagem bacteriana e resíduos de antibióticos. Emprega também regras para resfriamento do produto e coleta a granel [4]. De acordo com especificações presentes na Normativa 51 o produto pode ser definido dentro de algumas classes.

- Compreende-se como leite pasteurizado tipo A o leite produzido, beneficiado e envasado de acordo com as prescrições contidas na Normativa 51, em estabelecimento denominado granjas leiteiras. Esse tipo de leite pode ser classificado quanto ao teor de gordura em integral, padronizado, semi-desnatado ou desnatado. Deve ser obtido, obrigatoriamente através de equipamento de ordenha mecânica, pré-filtragem e bombeamento até o tanque de depósito em circuito fechado. Deve passar e ser aprovado por todos os testes bacteriológicos e de provas enzimáticas.
- Compreende-se, como leite cru refrigerado tipo B, o produto definido pela Normativa 51, integral quanto ao teor de gordura, produzido em estábulos leiteiros, refrigerado em propriedade rural produtora de leite e nela mantido pelo período máximo de 48 h, em temperatura igual ou inferior a 4°C, que deve ser atingida no máximo 3 h após o fim da ordenha, transportado para estabelecimento industrial, para ser processado, onde deve apresentar, no momento do seu recebimento, temperatura igual ou inferior a 7°C.
- Compreende-se, como leite tipo C, o produto definido pela Normativa 51 como aquele que não é submetido a qualquer tipo de tratamento térmico na fazenda leiteira, onde foi produzido, e integral quanto ao teor de gordura. Pode ser entregue até 10 h depois de sua obtenção, ou ser refrigerado a uma temperatura igual ou inferior a 4°C, sendo entregue, no máximo, em 24 h depois de sua estocagem.
- Compreende-se, como leite cru refrigerado, o leite independente do seu tipo, que é coletado na propriedade rural e transportado a granel, visando promover a redução geral de custos de obtenção e, principalmente, a conservação de sua qualidade até a recepção em estabelecimento submetido a inspeção sanitária oficial. O leite é refrigerado e mantido em temperaturas igual ou inferior a 4°C [4].

Depois de seguir as especificações de produção estabelecidas pela Normativa 51, o leite, ao chegar ao meio industrial, passa por um processo de filtragem para eliminar supostos detritos. Posteriormente, o leite resfriado é submetido a tratamento por calor, que tem como objetivo eliminar os germes. Este processo pode ser executado de diversas formas, como a pasteurização e a esterilização.

A pasteurização é realizada através do aquecimento do leite em determinada faixa de temperaturas, durante um tempo pré-determinado, seguido por um resfriamento imediato do leite.

A esterilização consiste no aquecimento do leite a altas temperaturas, com o objetivo de manter o produto em conservação durante um longo tempo e na temperatura ambiente [5]. Deve ser salientado que nenhum destes processos deve proporcionar alterações significativas em sua composição básica.

A ultrapasteurização é o processo utilizado para a obtenção do leite UHT (Ultra Alta Temperatura). Esse processo consiste no aquecimento do leite a 150°C por um tempo muito curto, de cerca de quatro segundos, seguido por um rápido resfriamento. Tanto os agentes causadores de doenças, que possam estar presentes no leite, quanto os não nocivos à saúde, são eliminados. Esta é a principal diferença entre o leite UHT longa vida e o leite pasteurizado.

A adulteração do leite ocorre geralmente, e com maior frequência, no campo ou meio rural, onde os produtores utilizam-se da adição de água ou soro do leite. Esse processo, de molhagem do leite, leva a uma diminuição na densidade e porcentagem de matéria gorda do leite. No meio industrial, o desnatamento ou acréscimo de leite desnatado é mais freqüente, e tem como consequência mais significativa, o aumento na densidade do leite, pois reduz a matéria gorda e uma diminuição na porcentagem de gordura do leite. Nota-se, então, a necessidade de avaliar a composição do leite tanto no meio rural como no meio industrial. No meio industrial a fiscalização do leite é obrigatória, mesmo após os processos de beneficiamento, para garantir que esses não modificaram sua composição.

2.3 O leite como líquido não-homogêneo

O leite é um líquido que tem todos seus componentes diluídos, com exceção da gordura, que tem menor densidade que a água. Por isso a gordura se encontra na forma de pequenas partículas que podem ter diferentes formatos e tamanho. Elas podem assumir desde a forma esférica, com diâmetro de 1 μm para o leite UHT, até formas não comuns e tamanhos da ordem de 10 μm para o leite *in natura*. Considerando o leite integral com 3,5% de gordura por volume, encontra-se cerca de 10^9 partículas por litro [10]. Estas partículas podem se encontrar de forma homogênea e com movimentos aleatórios ou de forma aglomerada em determinadas ocasiões.

A visualização da forma e a determinação do tamanho das partículas de gordura no leite são de grande importância, pois as medidas de atenuação e velocidade de propagação são muito sensíveis a estes parâmetros [10].

Utiliza-se o leite UHT pois é um líquido que apresenta características homogêneas, possui especificações vindas de fábrica, distribuição e formato de partículas regular. O leite UHT após o envase também necessita passar por inspeções para garantir a qualidade do produto, após os processos de beneficiamento.

2.4 Técnicas de análise de leite fluido

Estuda-se, neste ítem do trabalho, as principais formas de caracterização dos componentes do leite por técnicas físico-químicas, disponíveis no Laboratório de Biotecnologia da Faculdade de Engenharia-UNESP-Campus de Ilha Solteira, sob responsabilidade da professora Jacira dos Santos Isepon, para posterior confronto com técnicas estudadas por ultra-som. São abordados técnicas para medição de densidade, teor de gordura e quantidade de água adicionada ao leite.

2.4.1 Determinação da densidade

A densidade ρ de um corpo pode ser determinada de acordo com a relação entre sua massa m e o volume v , $\rho = m/v$. Para determinar a densidade, deve ser levada em conta a temperatura, pois o volume do líquido pode sofrer variação [5].

Para a determinação da densidade do leite utiliza-se um equipamento chamado de termolactodensímetro. Este aparelho é composto por uma esfera com flutuador e de uma coluna delgada, como apresentado na Figura 2.1, onde uma escala com graduações de 15°C a 45°C medem a variação de densidade de 1,015 a 1,045 kg/cm³.

Coloca-se o leite de forma a evitar a formação de espumas em uma proveta de

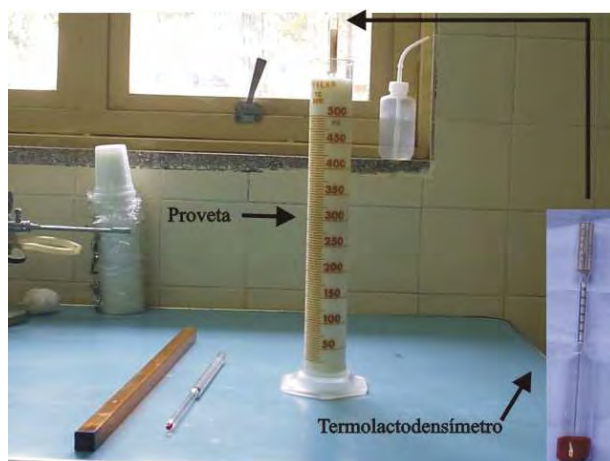


Figura 2.1. Esquema para medição da densidade: Proveta e Termolactodensímetro.

250 mL e mergulha-se o termolactodensímetro de tal forma que este possa flutuar livremente à altura do bordo superior, sem se fixar à parede da proveta. Para finalizar o processo deve ser feita a leitura na altura do nível do leite. Anota-se a temperatura, para posterior correção por uma tabela [5].

Outros métodos para determinação da densidade do leite com maior precisão, para pequenas amostras de leite, podem ser observadas na literatura. Um exemplo é a balança de Westfall [5]. Outro método conhecido utiliza o picnômetro para medição.

Para realizar a medição da densidade através do método do picnômetro, utiliza-se: balança de precisão, picnômetros, termômetro de precisão e amostras de água destilada e do líquido que se deseja medir a densidade.

Esse método consiste em utilizar como referência a densidade da água destilada, que é conhecida em diversas temperaturas. Assim, mede-se a massa de uma amostra de água, ou seja, mede-se a massa do picnômetro com e sem água e obtém-se a massa daquele determinado volume de água. Em seguida, mede-se a massa de uma amostra do líquido do qual deseja-se conhecer a densidade. O cálculo da densidade do líquido numa determinada temperatura é realizado através da Equação (2.1):

$$\rho_{amostra}(T) = \rho_{agua}(T) \frac{m_{amostra}(T)}{m_{agua}(T)}, \quad (2.1)$$

onde $\rho_{amostra}(T)$ é o valor da densidade da amostra em uma temperatura T , $\rho_{agua}(T)$ é o valor da densidade da água na temperatura T , $m_{amostra}(T)$ é a massa da amostra do líquido e $m_{agua}(T)$ é a massa do volume de água inserida no picnômetro.

2.4.2 Determinação de teor de gordura

Existe um grande interesse na obtenção da quantidade de gordura existente no leite, pois este é um dos componentes de maior importância comercial. Em condições normais o leite possui em média 3,5% de matéria gorda.

Existem diversos processos para a determinação de teor de gordura no leite. No Brasil, a técnica mais utilizada é o método ácido-butirômetro de Gerber. Entretanto, existe uma outra técnica chamada de método de Babcock que pode obter resultados com uma ordem de precisão maior [5]. O método de Gerber é baseado na propriedade que o ácido sulfúrico possui para dissolver a caseína do leite, sem atacar a matéria gorda, quando em concentração determinada [5]. Em outras palavras, ocorre a precipitação da caseína, seguida de sua destruição pelo H_2SO_4 , dosando-se em seguida em frascos especiais, a gordura flutuante.

Ao submeter o leite à centrifugação, a matéria gorda é separada dos outros componentes, os quais por serem mais leves se acumulam na parte superior do butirômetro, isto é, na haste graduada do mesmo. Na Figura, 2.2 observa-se a centrífuga de Gerber utilizada. A análise é realizada da seguinte forma:

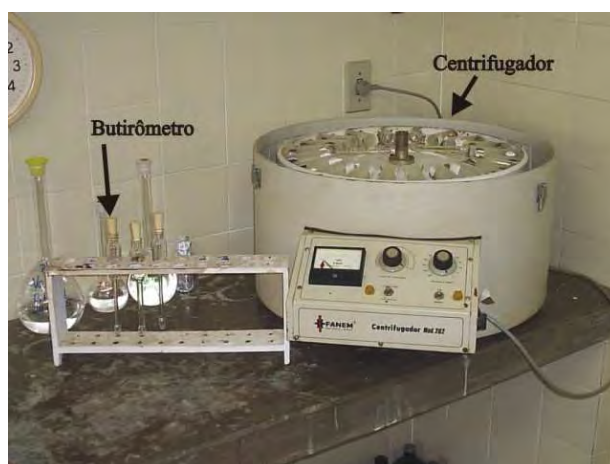


Figura 2.2. Centrífuga de Gerber e butirômetros.

- Coloca-se, no butirômetro de Gerber, 10 mL de H_2SO_4 com densidade (ρ) entre 1,820 a 1,825 kg/cm^3 .
- Adiciona-se 11 ml de leite, evitando que se misture ao ácido. Pode-se conseguir isto deixando que o leite escorra pelas paredes.
- Adiciona-se 1 ml de álcool amílico puro ($\rho = 0,815 \text{ kg}/\text{cm}^3$), com o objetivo de evitar a formação de espuma.
- O butirômetro deve ser fechado com uma rolha de borracha especial e coberto com um pano, na seqüência ele deve ser agitado para o H_2SO_4 decompor a caseína, obtendo-se um líquido homogêneo. Deve-se tomar cuidado, pois este processo forma uma reação exotérmica.
- Leva-se o butirômetro ao banho-maria por 3 a 5 minutos numa temperatura variando entre 65-70°C.
- Centrifuga-se o butirômetro, nesta temperatura, por 3 a 5 minutos a uma rotação entre 800 a 1200 por minuto.
- Logo em seguida, ajusta-se a rolha de borracha, para que a parte inferior da camada de gordura corresponda com o 0 (zero) da escala do butirômetro e faz-se a leitura. Cada divisão da escala equivale a 0,1% de matéria gorda do leite analisado.

O método de Gerber é preciso e bastante utilizado, mas demanda tempo elevado e manipulação de produtos perigosos para a saúde humana, como o ácido sulfúrico. Além da utilização de ácido, o método, exige pessoal qualificado para conduzir os experimentos. A reação entre os produtos causam aquecimento dos butirômetros, o que exige extremos cuidados em sua manipulação para evitar acidentes.

2.4.3 Determinação do ponto de congelamento do leite

Utiliza-se o método de crioscopia para medir o ponto de congelamento do leite, em comparação com os dissolventes puros. Este método utiliza um equipamento denominado crioscópio que é altamente sensível para medição de temperatura. Essa técnica visa detectar possíveis fraudes no leite por adição de água. Na Figura 2.3 apresenta-se o crioscópio.

Baseando-se na diferença do ponto de congelamento do leite e da água, o leite congela a $-0,55^{\circ}\text{C}$, com variações entre $-0,50^{\circ}\text{C}$ e $-0,57^{\circ}\text{C}$. A água congela a 0°C . Ao se adicionar uma quantidade de 1% de água o ponto de congelamento tende a subir aproximadamente $0,0055^{\circ}\text{C}$. Esta análise é feita da seguinte forma:

- Coloca-se no crioscópio água gelada, gelo em pedras e sal de cozinha, até que a temperatura no ambiente interior do aparelho fique entre $-5,0^{\circ}\text{C}$ e $-7,0^{\circ}\text{C}$.
- Determina-se o índice crioscópio de água destilada antes de iniciar as amostras, (C).
- Trabalha-se com uma amostra bem homogeneizada de leite, deixando resfriar até próximo de 0°C , para posteriormente colocá-la no crioscópio.
- Com o leite dentro do crioscópio deve-se observar a temperatura do mercúrio. Quando atingir $-0,55^{\circ}\text{C}$ deve-se colocar uma pedra de gelo na amostra, para ocorrer o choque térmico. Continua-se então movimentando a amostra até que o mercúrio suba à marca do ponto de congelamento. Deve-se tocar algumas vezes no termômetro para confirmar o ponto de congelamento (A).
- Determina-se o ponto de congelamento real da amostra (F):

$$F = A \pm C, \quad (2.2)$$



Figura 2.3. Crioscópio.

sendo as variáveis C e A , respectivamente, o ponto de congelamento da água e o ponto de congelamento da amostra. O índice N é um número da inspeção determinado por norma e seu valor é $-0,53$. Para se calcular a fraude por adição de água utiliza-se a Equação (2.3):

$$Fraude (\%) = \frac{(N - F)}{N} \times 100 \quad (2.3)$$

Na Tabela 2.2 apresenta-se o ponto de crioscopia versus a porcentagem de água, quando se conhece o índice crioscópico.

Tabela 2.2. Ponto crioscópico versus % de água.

Ponto de crioscópio	Quantidade de água adicionada
-0,53	3,63%
-0,52	5,45%
-0,51	7,27%
-0,50	9,09%
-0,46	16,36%
-0,45	18,18%
-0,44	20,00%
-0,43	21,81%
-0,42	23,63%
-0,37	32,72%
-0,36	34,54%
-0,28	50%

2.5 Comentários

Estudou-se as etapas de processamento do leite, onde se destaca as formas com que o leite é tratado ao chegar no meio industrial. Assim, através das especificações de processos de produção, classifica-se o leite bovino em suas categorias distintas.

Destacou-se também as técnicas de medição convencionais. Estudou-se aqui as técnicas para caracterização das amostras do leite, disponíveis no laboratório de biotecnologia da FE-UNESP-Campus de Ilha Solteira. O estudo dessas técnicas foi de suma importância para se adquirir certa habilidade no manuseio dos equipamentos. Lembrando que, em primeira instância, as técnicas aqui estudadas foram utilizadas para se obter as características do leite para que se possa relacioná-las com os parâmetros acústicos.

A análise realizada para densidade, pelo método que utiliza o termolactodensímetro pode ser precisa, mas não possui uma boa resolução. Já o método que utiliza o picnômetro é preciso e, com controle de temperatura, pode-se obter uma alta resolução.

A medição de teor de gordura pelo método de Gerber possui alta precisão e resolução, mas utiliza ácido sulfúrico que pode apresentar certo perigo de manipulação.

Foram realizadas visitas técnicas a laticínios, onde foi observada a importância das análises para constatação da qualidade do produto captado e a necessidade do desenvolvimento de novas técnicas de análise para melhorar ou torná-las mais rápidas.

Capítulo 3

Técnicas de ultra-som e célula de medição

Neste capítulo serão apresentadas as técnicas de ultra-som utilizadas para caracterização de líquidos. Serão abordados os princípios de propagação de ondas em meios não-homogêneos e as características e o funcionamento da célula de medição de parâmetros acústicos utilizada [14].

3.1 Parâmetros acústicos

A caracterização de materiais por ultra-som pode ser feita com base na medição de parâmetros da onda acústica como a velocidade de propagação, fase e amplitude.

Os parâmetros mais utilizados para a caracterização de líquidos por ultra-som são a velocidade de propagação, o coeficiente de atenuação e a densidade. Esses parâmetros estão relacionados com as propriedades físico-químicas dos materiais envolvidos, tal como sua composição, estrutura molecular e estado físico. Esta relação pode ser estabelecida por meio de experimentos realizados em laboratório, pela calibração de curvas características que relacionam as propriedades de interesse do líquido com as propriedades acústicas medidas ou através de modelos matemáticos que descrevem a propagação das ondas de ultra-som pelos materiais. Nos próximos itens serão descritos de forma simplificada os parâmetros: velocidade de propagação, atenuação e densidade.

3.1.1 Velocidade de propagação

A velocidade de propagação do som depende da estrutura do material no qual a onda se propaga e de sua forma ou modo de propagação. Em líquidos, em geral, são utilizadas ondas longitudinais ou de compressão, em que as partículas do meio movimentam-se na mesma direção da propagação da onda [23].

Para os fluidos, de forma simplificada, a velocidade (c) depende do módulo de elasticidade adiabático (β) e da densidade (ρ). A Equação (3.1) representa esta relação para ondas de compressão:

$$c = \sqrt{\frac{\beta}{\rho}}. \quad (3.1)$$

O módulo adiabático e a densidade dos materiais dependem de sua estrutura, da composição e do estado físico, portanto a velocidade de ultra-som pode ser utilizada para fornecer informações sobre as propriedades dos materiais. Existem dois métodos bem explorados para se obter a velocidade. O primeiro é através do conhecimento do comprimento de onda (λ) a uma determinada frequência (f), tendo $c = \lambda f$ e o outro por meio da obtenção do tempo de propagação (Δt) da onda por uma determinada distância (L), $c = L/\Delta t$.

3.1.2 Coeficiente de atenuação

O coeficiente de atenuação pode ser definido como a perda de energia mecânica durante a propagação de uma onda acústica num meio, causada por uma variedade de mecanismos físicos. Esse fenômeno causa uma diminuição na amplitude das ondas ultra-sônicas ao se propagar através do material [23].

A maior causa da atenuação são a absorção e o espalhamento. A absorção é observada em meios homogêneos, que ocorre devido a mecanismos físicos que convertem energia acústica armazenada em energia térmica [24]. O espalhamento ocorre em meios heterogêneos, isto é, em materiais homogêneos, mas que contém partículas de outros materiais, tais como emulsões, suspensões e espumas. O processo ocorre quando uma onda de ultra-som incide numa descontinuidade e é espalhada em outras direções, ou seja, devido a mecanismos que defletem ou espalham a energia da onda [25].

Atenuação em meios homogêneos

Os principais mecanismos causadores de atenuação acústica em meios homogêneos são: a viscosidade do meio, as perdas termodinâmicas e as perdas devido à relaxação.

As perdas causadas pela viscosidade são devido aos movimentos relativos entre as partes circunvizinhas, durante a compressão e rarefação que caracterizam a propagação da onda [24, 26].

As perdas térmicas ocorrem devido ao gradiente de temperatura no meio quando ele é comprimido e expandido. A condutividade térmica do meio faz com que ocorram trocas térmicas no processo, sendo assim essa condutividade contribui para a perda de energia térmica [23, 24, 26].

As perdas devido à relaxação ocorrem a partir das trocas de energia, que estão relacionadas à conversão da energia cinética das moléculas, translacional, em:

- energia potencial, com um novo rearranjo estrutural das moléculas;
- energia interna vibracional e ou rotacional das moléculas.

Quando uma onda acústica propaga-se, ocorre uma reorganização das moléculas em grupos de tamanho ou formatos diferentes. Após a passagem da onda acústica, as moléculas voltam a organizar-se como antes, e então ocorre a relaxação. A relaxação térmica ocorre devido à conversão de energia de translação em energia interna, ou seja, em energia associada à rotação e vibração das moléculas [23].

Atenuação em meios não-homogêneos

Um meio não-homogêneo pode ser representado por um meio homogêneo contendo partículas em suspensão. Para estes meios, além das perdas devido ao meio, adicionam-se as perdas devido às partículas em suspensão.

Existem basicamente quatro mecanismos, de grande importância, causadores de perdas em meios não-homogêneos [15, 24, 26]:

- Perda viscosa;
- Perda térmica;
- Perda por espalhamento;
- Perdas intrínsecas.

As perdas causadas pela viscosidade ocorrem devido a uma onda de cisalhamento gerada pelos movimentos relativos entre as partículas suspensas no meio. Estes movimentos, de translação e pulsação radial, são causados devido à diferença entre as densidades da partícula e do meio fluido [24, 26].

As perdas térmicas são causadas pela geração de um gradiente de temperatura na região próxima à superfície da partícula. Este gradiente de temperatura aparece devido à interação entre a pressão e a temperatura [24, 26].

As perdas causadas por espalhamento são referentes ao redirecionamento do fluxo de energia da onda devido à presença de partículas suspensas no meio, tendo parte da energia desviada e não absorvida pelo receptor. As perdas por espalhamento acústico não produzem dissipação de energia, sendo diferentes dos mecanismos de perdas térmica e viscosa. Essas perdas são geométricas e significativas apenas para partículas grandes, acima de $1\ \mu\text{m}$, em altas frequências, acima de 50 MHz [25].

Os mecanismos de perda intrínsecos ocorrem devido à interação da onda acústica com materiais das partículas e do meio como se fossem meios homogêneos [24, 26].

3.1.3 Impedância acústica e densidade

A impedância acústica de um material relaciona-se com sua propriedade de transmissão de uma vibração acústica, com maior ou menor facilidade.

A pressão acústica de um material p , é definida como a diferença entre a pressão instantânea em qualquer ponto do material e a pressão de equilíbrio deste, ou seja, é a pressão em excesso neste ponto. Ao receber estímulos, as partículas de um material vibram com determinada velocidade, u , que depende da pressão acústica e da dificuldade do meio em transmitir tais vibrações [23].

Define-se impedância acústica específica, Z , como a razão da pressão acústica em um meio pela velocidade da partícula associada, onde o sinal depende do sentido de propagação da onda [23].

$$Z = \pm \frac{p}{u}, \quad (3.2)$$

Considerando ondas planas, define-se a impedância acústica característica como o produto da densidade de equilíbrio do material, ρ_o , pela velocidade de propagação do som neste material, c [23]:

$$Z = \pm \rho_o c. \quad (3.3)$$

3.1.4 Fenômenos de transmissão

Quando uma onda de pressão incidente com magnitude p_i se propaga num meio 1 de impedância acústica característica Z_1 e encontra uma interface com outro meio 2, com impedância acústica diferente Z_2 , parte dessa onda é refletida, com magnitude p_r (onda refletida), e parte é transmitida com pressão p_t (onda transmitida) como pode ser visto na Figura 3.1. Esta análise pode ser simplificada quando se assume a incidência normal, de uma onda plana longitudinal em uma interface plana entre dois meios.

Seja uma onda plana longitudinal se propagando na direção $\mathbf{x}$, que encontra uma interface em $x = 0$, entre o meio 1 e um meio 2, os coeficientes de reflexão e transmissão em uma interface são definidos como a razão entre a pressão acústica da onda refletida e transmitida, respectivamente, pela pressão incidente, como mostrado na Equação (3.4).

$$R = \frac{p_r}{p_i}, \quad T = \frac{p_t}{p_i}. \quad (3.4)$$

A partir das condições de contorno em $x = 0$ para pressão e velocidade da partícula, obtêm-se as Equações (3.5) e (3.6) [23].

$$p_i + p_r = p_t \quad (3.5)$$

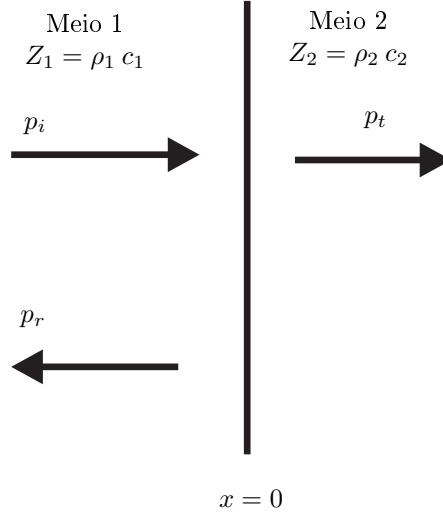


Figura 3.1. Incidência normal de uma onda a uma interface.

e

$$u_i + u_r = u_t. \quad (3.6)$$

Dividindo a Equação (3.5) pela Equação (3.6), obtém-se:

$$\frac{p_i + p_r}{u_i + u_r} = \frac{p_t}{u_t}. \quad (3.7)$$

Sabendo-se que as ondas incidentes e refletidas encontram-se no meio 1, assim como a onda transmitida encontra-se no meio 2, utiliza-se a Equação (3.2) para obter uma equação com incidência normal (3.8).

$$Z_2 = Z_1 \frac{p_i + p_r}{p_i - p_r}. \quad (3.8)$$

Dividindo-se o numerador e o denominador da Equação (3.8) por p_i , obtém-se:

$$Z_2 = Z_1 \frac{1 + R}{1 - R}. \quad (3.9)$$

Ao isolar a variável R na Equação (3.9) obtém-se o coeficiente de reflexão na interface entre os meios 1 e 2

$$R = \frac{Z_2 - Z_1}{Z_1 + Z_2}. \quad (3.10)$$

Dividindo-se a Equação (3.5), em ambos os lados por p_i , obtém-se: (3.11).

$$1 + R = T. \quad (3.11)$$

Substituindo a Equação (3.11) no resultado obtido para o coeficiente de reflexão de uma interface (3.10), obtém-se uma nova fórmula para o coeficiente de transmissão (3.12):

$$T = \frac{2Z_2}{Z_1 + Z_2}. \quad (3.12)$$

Por meio da Equação (3.9), pode-se calcular algumas variáveis importantíssimas relacionadas com este trabalho. Relacionando-se a fórmula (3.2) com (3.9), obtém-se a Equação para a densidade (3.13).

$$\rho_2 = \frac{\rho_1 c_1}{c_2} \frac{1 + R}{1 - R}. \quad (3.13)$$

Assim, tendo-se um meio 1 conhecido (de referência), obtém-se os valores de ρ_1 e c_1 , através da determinação do coeficiente de reflexão e da velocidade de propagação c_2 , pode se obter a densidade do meio 2.

3.2 Célula de medição

A célula de medição por ultra-som é formada por três componentes: transdutor duplo-elemento (TDE), câmara de amostra e refletor. Na Figura 3.2 apresenta-se o esquema da célula. A célula foi inicialmente concebida por Adamowski [14] para medição de densidade de líquidos por ultra-som.

A célula de medição de líquidos utiliza um transdutor duplo-elemento, que é formado por um emissor e um receptor de ultra-som. O receptor utilizado é constituído de um material de P(VDF-TrFE) (Polyvinylidene Fluoride Tri-Fluoro-Ethylene) e o emissor cerâmico comercial. O receptor foi modelado de forma a evitar e/ou eliminar o efeito de difração acústica, que possui uma área suficientemente grande para que

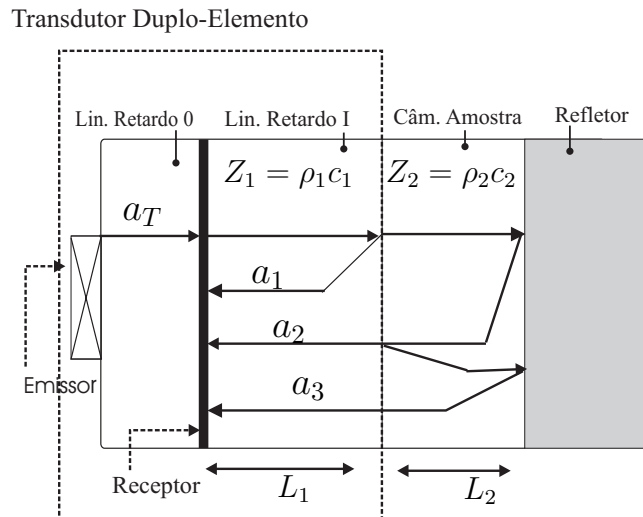


Figura 3.2. Célula de medição.

possa interceptar todo o campo acústico. Esta membrana piezelétrica é metalizada em ambas as faces. Nas Figuras 3.3 e 3.4 apresentam-se as fotos da célula utilizada.

O receptor e o emissor são separados por um material sólido homogêneo para que se possa trabalhar de forma individual os sinais elétricos da emissão do pulso e da recepção dos ecos. Este meio homogêneo que os separa foi chamado de linha de retardo, essas linhas de retardo são confeccionadas com material que possui propriedades conhecidas (densidade e comprimento) no caso, acrílico.

A membrana de PVDF possui uma espessura de $52\text{ }\mu\text{m}$, e seu diâmetro foi escolhido para que possa interceptar todo o campo acústico gerado pelo transdutor e pelos ecos recebidos, na faixa de frequência de trabalho. O diâmetro da linha de retardo deve ter tamanho tal que a abertura do feixe acústico não cause reflexões nas paredes laterais. O comprimento foi calculado de forma que não ocorra superposição

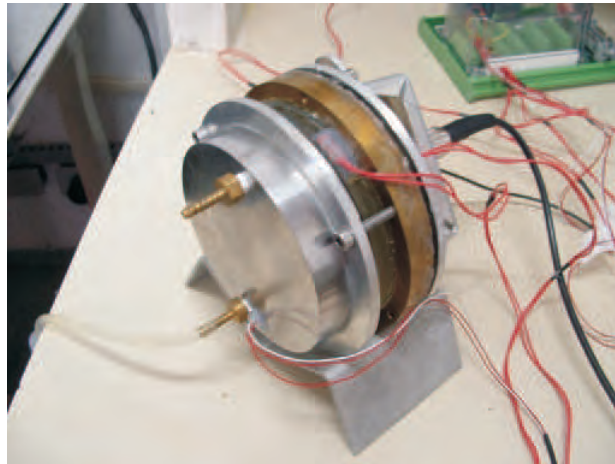


Figura 3.3. Foto da célula de medição: vista superior.

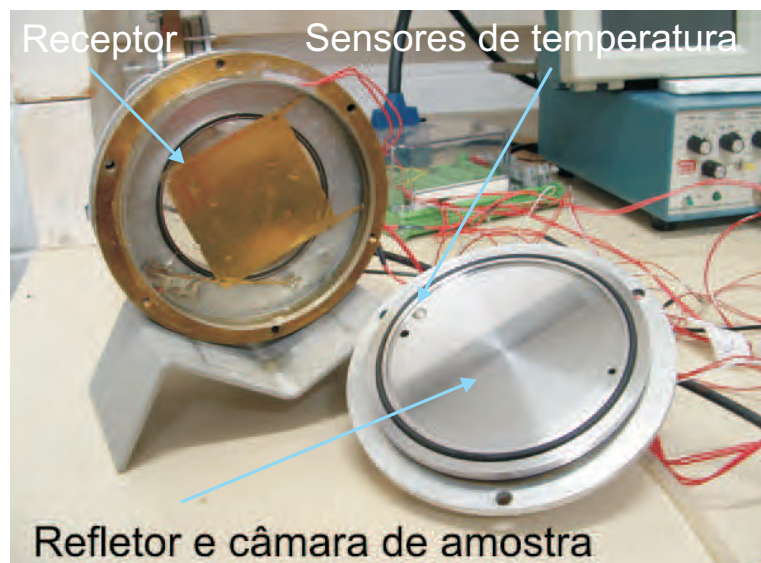


Figura 3.4. Foto da célula de medição: TDE à esquerda e à direita, refletor e câmara de amostra.

dos sinais recebidos. Para que não ocorra a geração de pulsos espúrios e se tenha incidência normal, as faces das linhas de retardo devem ser confeccionadas de tal forma a manter um paralelismo entre as interfaces [14].

A membrana receptora é separada do líquido pela linha de retardo I que constitui o meio 1. O meio 2 é o líquido depositado na câmara de amostra, o qual se deseja medir suas propriedades. O meio 3 é o refletor, o qual é utilizado para reflexões dos sinais de ultra-som.

O refletor é constituído de material de inox, um material de impedância acústica elevada. O fato de possuir características acústicas conhecidas e ser um material de impedância acústica maior, se comparado à impedância acústica da amostra, o torna um bom refletor tendo uma reflexão quase total da onda, aproximadamente 94% [10].

O transdutor emissor gera um pulso ultra-sônico que se propaga pela linha de retardo 0 até alcançar o receptor de PVDF. Assim, o receptor é sensibilizado podendo medir o sinal transmitido $a_T(t)$. O sinal transmitido propaga-se pela linha de retardo I até alcançar a interface com a amostra, onde de forma parcial é refletido e transmitido. A onda é refletida e sensibiliza o receptor (a_1), enquanto a onda transmitida propaga-se na amostra, e é refletida pelo refletor, retorna e novamente encontra-se com a interface da amostra. Da mesma forma, o sinal é parcialmente refletido e transmitido, sendo que neste instante a onda transmitida sensibilizará o receptor (a_2). O sinal refletido irá continuar o caminho entre o refletor e a interface com a amostra sempre transmitindo um sinal na mesma. Nas figuras 3.5 e 3.6 apresentam-se um exemplo do sinal transmitido e os três ecos recebidos, nos domínios do tempo e da frequência, respectivamente. Através deste procedimento são calculados os parâmetros acústicos em questão: densidade, velocidade de propagação e coeficiente de atenuação do líquido [14].

3.2.1 Coeficiente de reflexão

O coeficiente de reflexão na interface linha de retardo I/ líquido (R_{12}), é obtido em função dos sinais de eco $a_1(t)$, $a_2(t)$, $a_3(t)$. Este coeficiente de reflexão depende das impedâncias acústicas dos meios 1 e 2, sendo calculado a partir da Equação (3.14):

$$R_{12} = \frac{Z_2 - Z_1}{Z_2 + Z_1} = \left(\frac{1}{1 - x} \right)^{1/2}. \quad (3.14)$$

A variável x é determinada a partir dos espectros dos sinais de eco $a_1(t)$, $a_2(t)$, $a_3(t)$, através de duas metodologias destacadas a seguir. Dois métodos utilizados para o cálculo do coeficiente de reflexão são: o método da frequência [14] e de energia [13].

O método da frequência consiste em calcular o coeficiente de reflexão a partir dos espectros de magnitude dos ecos $a_1(t)$, $a_2(t)$, $a_3(t)$ em uma frequência determinada

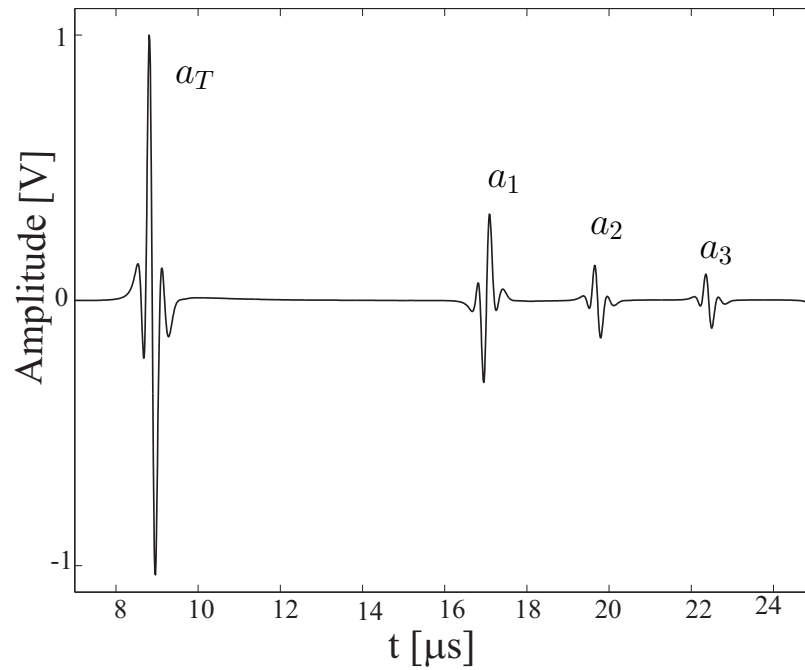


Figura 3.5. Gráfico dos sinais de eco no domínio do tempo.

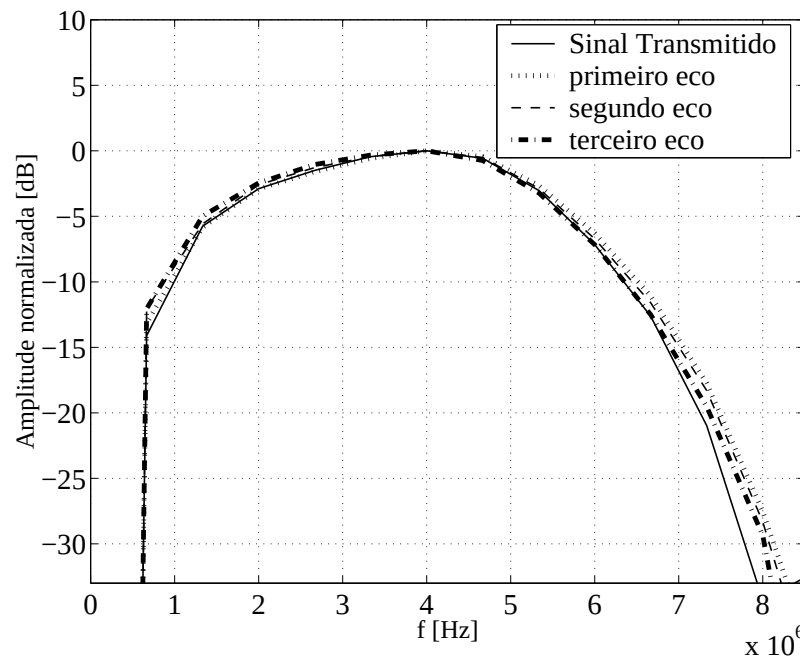


Figura 3.6. Gráfico dos sinais recebidos no domínio da frequência.

[14].

Os sinais de eco no domínio da frequência são escritos em função do eco transmitido $a_T(t)$ [13]:

$$A_1(f) = A_T(f) R_{12} e^{-2\alpha_1(f) l_1} e^{-j2\pi f \Delta t_1}, \quad (3.15)$$

$$A_2(f) = A_T(f) T_{12} R_{23} T_{21} e^{-2\alpha_1(f) l_1} e^{-2\alpha_2(f) l_2} e^{-j2\pi f (\Delta t_1 + \Delta t_2)}, \quad (3.16)$$

$$A_3(f) = A_T(f) T_{12} R_{23}^2 R_{21} T_{21} e^{-2\alpha_1(f)l_1} e^{-4\alpha_2(f)l_2} e^{-j2\pi f(\Delta t_1 + 2\Delta t_2)} . \quad (3.17)$$

Tem-se que R_{jk} e T_{jk} , são respectivamente, os coeficientes de reflexão e transmissão nas interfaces entre os meios j e k , para $j, k = 1, 2$ ou 3 , α_i o coeficiente de atenuação, l_i o comprimento e o Δt_i é o tempo que a onda leva para propagar-se duas vezes pelo meio de comprimento l_i , onde $i=1,2,3$. O coeficiente de reflexão R_{12} , é calculado reagrupando as equações das amplitudes, e x uma variável obtida a partir do reagrupamento das equações dos espectros de magnitude dos ecos $a_1(t)$, $a_2(t)$, $a_3(t)$, que pode ser determinada em função de uma frequência f_i determinada, através da Equação (3.18) [13]:

$$|x| = \frac{|A_2(f)|^2}{|A_1(f)| |A_3(f)|} \Big|_{f=f_i} . \quad (3.18)$$

O método da energia consiste no cálculo da energia dos sinais dos ecos. Ao contrário do método da frequência, cujo cálculo é realizado numa frequência determinada, no método de energia ocorre o cálculo numa faixa de frequências.

Segundo Higuti [13], a Equação (3.18), escrita em função da densidade espectral de energia dos sinais em questão: $a_1(t)$, $a_2(t)$, $a_3(t)$, pode ser tomada de forma alternativa sendo descrita de acordo com a Equação (3.19).

$$|x_E| = \frac{\int_{-\infty}^{\infty} |A_2(f)|^2 df}{\int_{-\infty}^{\infty} |A_1(f)| |A_3(f)| df} . \quad (3.19)$$

Na dificuldade em calcular ruídos numa determinada frequência, o método de energia, segundo Higuti [13] possui melhores resultados para o cálculo do coeficiente de reflexão, por fazer uma média do ruído na frequência, resultando em desvios menores.

3.2.2 Velocidade de propagação

A velocidade de propagação da onda acústica no meio líquido em questão é calculada a partir da medição da diferença de tempo de propagação entre dois ecos consecutivos, $a_1(t)$ e $a_2(t)$ ou $a_2(t)$ e $a_3(t)$. A Equação (3.20) é utilizada para o cálculo da velocidade na amostra.

$$c_2 = \frac{2L_2}{\Delta t_{12}} , \quad (3.20)$$

sendo a variável L_2 o comprimento da câmara de amostra do líquido, conhecida através de um processo de calibração que utiliza água destilada. A calibração consiste em calcular o comprimento L_2 da câmara de amostra, a partir do cálculo de Δt e da velocidade de propagação c da água destilada, que é tabelada em função da temperatura.

O tempo de propagação entre os sinais, Δt_{12} , é calculado através da técnica de correlação cruzada. Esta técnica é bastante utilizada devido à precisão, na qual calcula o intervalo de tempo entre sinais. Ela faz a determinação deste tempo relacionando o instante de máximo da função de correlação entre dois sinais consecutivos. A função de correlação entre dois sinais de tempo discreto $a_i[n]$ e $a_j[n]$ pode ser calculada pela Equação (3.21).

$$C[m] = \sum_{n=-\infty}^{\infty} a_i[n]a_j[n+m] , \quad (3.21)$$

os índices i e j são referentes aos sinais de eco em questão, sendo i igual a 1 e j igual a 2 ou i igual a 2 e j igual a 3.

3.2.3 Densidade e coeficiente de atenuação

A densidade da amostra é calculada a partir dos valores R_{12} e c_2 , coeficiente de reflexão e velocidade de propagação da onda, obtidos a partir das Equações (3.14) e (3.20) respectivamente. Como mencionado anteriormente, esses parâmetros são calculados a partir dos sinais $a_1(t)$, $a_2(t)$ e $a_3(t)$. Através da Equação (3.13) determina-se a densidade da amostra, ρ_2 [13].

$$\rho_2 = \frac{\rho_1 c_1 (1 + R_{12})}{c_2 (1 - R_{12})} , \quad (3.22)$$

sendo que ρ_1 é o valor da densidade da linha de retardo I, que deve ser conhecido, c_1 é a velocidade de propagação na linha de retardo I, adquirida a partir dos sinais $a_T(t)$ e $a_1(t)$, de forma equivalente à Equação (3.20) [13].

Outra variável importante para o processo de caracterização das amostras, é o coeficiente de atenuação da amostra. Esta é utilizada em inúmeras aplicações para caracterização de emulsões e suspensões.

O coeficiente de atenuação é calculado a partir do coeficiente de reflexão R_{23} , adquirido da interface líquido/refletor e das magnitudes das transformada de Fourier dos sinais ($|A_1(f)|$) e ($|A_2(f)|$). O coeficiente de reflexão R_{23} é obtido a partir da Equação (3.23):

$$R_{23} = \frac{Z_3 - Z_2}{Z_3 + Z_2} . \quad (3.23)$$

Assumindo-se que as propriedades acústicas do refletor sejam conhecidas e variem pouco com a temperatura de trabalho, o coeficiente de atenuação é determinado a partir da Equação (3.24) relacionando as Equações (3.15) e (3.16) ou (3.16) e (3.17) [14].

$$\alpha_2(f) = \frac{1}{2L_2} \ln \left[\frac{|A_1(f)|}{|A_2(f)|} \frac{(1 - R_{12}^2)}{R_{12}} R_{23} \right] . \quad (3.24)$$

O coeficiente de atenuação é utilizado para caracterização de gordura em amostras de leite [25]. Um problema previsto para este tipo de amostra é a baixa atenuação. Este problema é devido ao pequeno comprimento da câmara de amostra utilizado na célula em questão, proporcionando uma pequena atenuação no sinal.

3.3 Comentários

Neste capítulo foi realizado um estudo inicial sobre os fenômenos de propagação e transmissão de ondas acústicas em meios homogêneos e não-homogêneos.

O capítulo também apresentou as principais características de funcionamento da célula de medição utilizada neste trabalho e os parâmetros acústicos a serem medidos foram apresentados. Apresentou-se também a metodologia utilizada para medição dos parâmetros acústicos. A utilização de forma correta da célula é de extrema importância para que não se cometam erros na medição de parâmetros, por isso todo o princípio de funcionamento da célula deve ser estudado e entendido.

As técnicas apresentadas nesse capítulo foram utilizadas para a obtenção dos resultados experimentais para caracterização de leite. Esses resultados são apresentados no capítulo 5.

Capítulo 4

Redes neurais artificiais

Neste capítulo serão apresentados conceitos sobre Redes Neurais Artificiais, suas vantagens, funcionamento, implementação e eficiência. Também serão abordados assuntos como: neurônios, comparando neurônios biológicos e neurônios artificiais; funções de ativação de neurônios artificiais; aprendizagem do neurônio e redes do tipo Perceptron “(*MPL-Multi Layer Perceptron*)” e Adaline. Os aspectos básicos apresentados são importantes como documentação, por ser tratar do primeiro trabalho do grupo na área de redes neurais.

4.1 Redes neurais artificiais (RNAs)

Segundo Braga [16], Redes Neurais Artificiais são sistemas paralelos distribuídos compostos por unidades de processamento simples, chamados neurônios, que calculam determinadas funções matemáticas, normalmente não-lineares. Tais unidades são dispostas em uma ou mais camadas e interligadas por um grande número de conexões, geralmente unidirecionais. Na maioria dos modelos estas conexões estão associadas a pesos, os quais são utilizados para armazenar o conhecimento representado no modelo e para ponderar a entrada recebida por cada neurônio da rede. Redes Neurais possuem tendência natural para armazenar conhecimentos experimentais e torná-los disponíveis para o uso, que se assemelha ao cérebro humano em dois aspectos fundamentais:

- As conexões entre os neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido.
- O conhecimento é adquirido pela rede por meio de dados do ambiente, num processo de aprendizagem. O processo de treinamento é chamado de “Algoritmo de Aprendizagem”, que tem como finalidade ajustar os pesos sinápticos da rede de uma forma ordenada para alcançar um objetivo desejado.

Resolver problemas utilizando-se RNAs é uma excelente alternativa, uma vez que o processamento é estruturalmente paralelo e simultâneo, garantindo um desempenho superior ao de outros métodos como por exemplo, árvore de decisões e outros [27].

Antes de utilizar uma RNA, um processo de treinamento é necessário para minimizar os erros que eventualmente irão acontecer e, conseqüentemente, deixá-la apta para resolver problemas, extraíndo informações dos exemplos apresentados à rede, e posteriormente tornando-se capaz de gerar soluções para seus problemas.

Existem vários métodos de aprendizagem, mas o que é importante ressaltar é a capacidade de aprender através de conjuntos de dados (exemplos) apresentados à rede e a de generalizar a informação assimilada, a fim de solucionar problemas semelhantes, mas que ainda não são conhecidos pela RNA. Essas duas características são os principais diferenciais em relação a outros modelos, e o que torna esse método atrativo e eficiente para solução de problemas complexos.

4.2 Neurônio biológico

Um neurônio biológico possui três partes distintas: o corpo celular, os dendritos e o axônio. Os dendritos têm como função receber e conduzir os impulsos nervosos (informações) até o corpo celular do neurônio, onde essa informação é processada, e outros impulsos são gerados. Esses novos impulsos são transmitidos através do axônio até os dendritos de outros neurônios. O ponto de contato entre a terminação axônica de um neurônio e o dendrito de outro é chamado de sinapse. É pela sinapse que os nodos se unem funcionalmente, formando redes neurais. Essas sinapses funcionam como válvulas que controlam as transmissões de impulsos, ou seja, o fluxo de informação entre os nodos na rede neural. Se o percentual de impulsos recebidos, em um intervalo curto de tempo, é suficientemente alto, a célula é ativada, produzindo um impulso que é transmitido para as células seguintes. É assim que ocorre a maioria das operações realizadas pelo cérebro humano [16].

4.3 Neurônio artificial de McCulloch e Pitts

O primeiro modelo artificial de que se tem notícia foi o proposto por Warren McCulloch e Walter Pitts [28]. Esse modelo foi inspirado no conhecimento que se tinha sobre neurônios biológicos, em 1943. A descrição matemática desse modelo resultou em um modelo com n terminais de entrada $x_1, x_2, \dots, x_n$ (dendritos) e apenas um terminal de saída y (axônio). Para emular o comportamento das sinapses, os terminais de entrada do neurônio têm pesos acoplados $w_1, w_2, \dots, w_n$ cujos valores podem ser positivos ou negativos, dependendo das sinapses correspondentes serem inibitórias ou excitatórias. O efeito de uma sinapse particular i no neurônio pós-

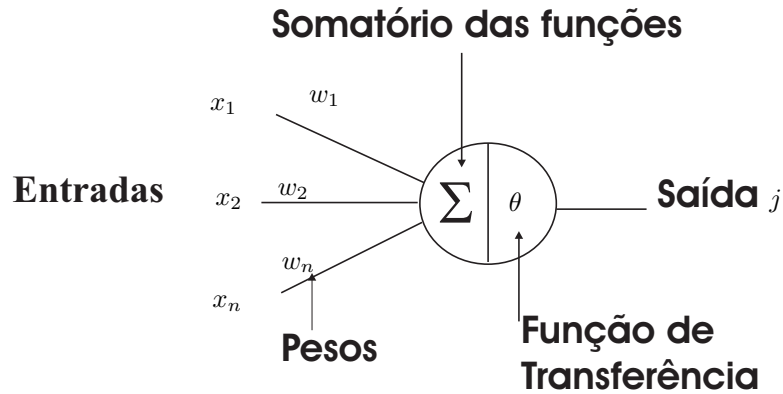


Figura 4.1. Neurônio artificial de McCulloch e Pitts.

sináptico é dado por $x_i w_i$. Uma representação gráfica desse modelo está indicada na Figura 4.1.

A ativação do nodo MCP (McCulloch e Pitts), se dá através da soma ponderada dos produtos entre n terminais de entrada e os pesos associados a cada entrada, e essa soma determina se o neurônio deve ou não ativar. Para que o neurônio se ative é preciso que a soma ponderada satisfaça a função de ativação. Na descrição original do modelo MCP, a função de ativação é dada pela função de limiar escrita na Equação (4.1). O nodo MCP terá sua saída ativa quando:

$$\sum_{i=1}^n x_i w_i \geq \theta, \quad (4.1)$$

sendo n o número de entradas do neurônio, w_i é o peso associado à entrada x_i e θ é o limiar (*threshold*) do neurônio.

4.4 Funções de ativação

A partir do modelo de McCulloch e Pitts, foram derivados outros modelos com diferentes funções de ativação: função linear, rampa, degrau e sigmoidal [16].

Na função de ativação linear $y = a.x$, sendo a um número real que define a saída linear para os valores de entrada x , y é a saída. Na Figura 4.2, apresenta-se a função de ativação linear.

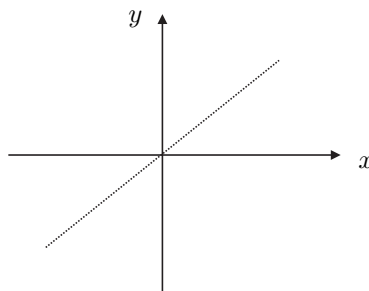


Figura 4.2. Função de ativação linear.

A função linear pode ser restringida para produzir resultados na faixa $[-\gamma, +\gamma]$, e neste caso ela passa a ser uma função rampa. Na Figura 4.3 é apresentada a função rampa.

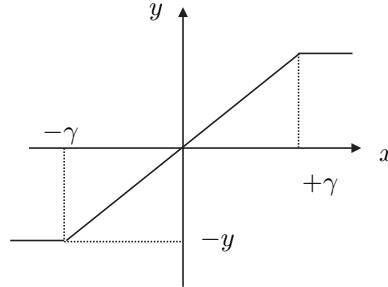


Figura 4.3. Função de ativação rampa.

$$y = \begin{cases} +\gamma, & \text{se } x \geq +\gamma \\ x & \text{se } |x| < +\gamma \\ -\gamma & \text{se } x \leq -\gamma \end{cases}.$$

Na Figura 4.4 apresenta-se a função degrau que é simples de se implementar. Se x for maior que 0 a saída será $+\gamma$, caso contrário será $-\gamma$:

$$y = \begin{cases} +\gamma, & \text{se } x > 0 \\ -\gamma & \text{se } x \leq 0 \end{cases}.$$

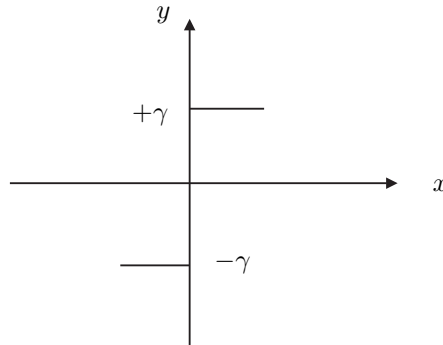


Figura 4.4. Função degrau.

A função sigmoideal, é uma função semilinear e limitada. As funções sigmoideais são encontradas em projetos de diversos modelos nas mais variadas áreas. Na Figura 4.5 apresenta-se uma função logística sigmoideal. Uma das funções sigmoideais mais utilizadas é definida pela Equação (4.2):

$$y = \frac{1}{1 + e^{x/T}}, \quad (4.2)$$

sendo o parâmetro T responsável pela definição da suavidade da curva, apresentada na Figura 4.5.

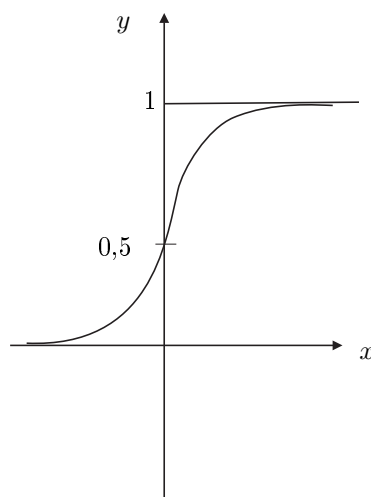


Figura 4.5. Função de ativação sigmoideal.

4.5 Aprendizagem

Toda rede neural passa por um processo de aprendizagem antes de começar a solucionar qualquer problema. Durante esse processo a rede neural extrai informações relevantes de padrões de informação apresentados a ela, criando assim uma representação própria para o problema. Isso é um processo iterativo de ajuste de parâmetros da rede, mais especificamente dos pesos das conexões entre as unidades de processamento (neurônios), que armazenam o conhecimento durante a aprendizagem.

Existem diversos métodos de aprendizagem, cada um com suas vantagens e desvantagens, mas a sua essência e a grande diferença continuam sendo os pesos e a forma como eles são modificados.

4.5.1 Aprendizado supervisionado

Segundo Braga [16], o aprendizado supervisionado consiste em fornecer à rede um conjunto de dados de entrada e o valor das saídas correspondentes. Com base nos valores de entrada, a RNA calcula o valor de saída. Caso esse valor seja diferente do valor desejado é calculado o erro (diferença) entre a saída obtida e a saída desejada. Com base nesse erro, a RNA realiza atualizações nos pesos dos neurônios com o objetivo de que a saída obtida pela rede se aproxime ao máximo da saída desejada.

A grande desvantagem desse método é que, na ausência do supervisor, a rede neural não irá aprender novas estratégias, ou seja, seus pesos não serão atualizados e assim sendo novas situações que não são cobertas pelos conjuntos de dados apresentados no treinamento não terão uma solução adequada. Na Figura 4.6, apresenta-se o diagrama de blocos do aprendizado supervisionado.

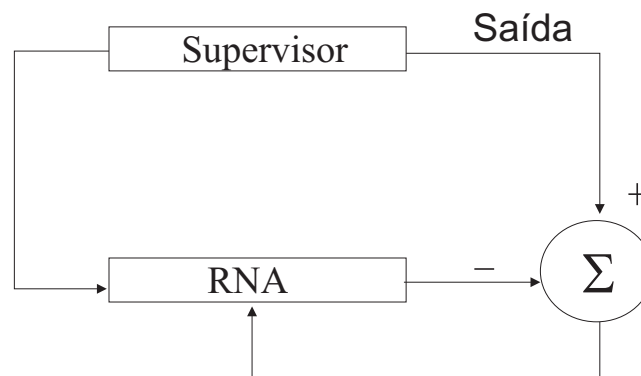


Figura 4.6. Aprendizado supervisionado.

4.5.2 Aprendizado não-supervisionado

O aprendizado não-supervisionado sugere que os algoritmos não possuem um supervisor, mas somente os dados de entrada estão disponíveis. Desta forma não há exemplos rotulados da função a serem aprendidos pela rede. Quando os parâmetros livres da rede tiverem se ajustado às regularidades estatísticas dos dados de entrada, ela desenvolve a habilidade de formar representações internas para codificar características da entrada, e desse modo criar automaticamente novas classes. Na Figura 4.7 é apresentado o diagrama de blocos do aprendizado não-supervisionado.

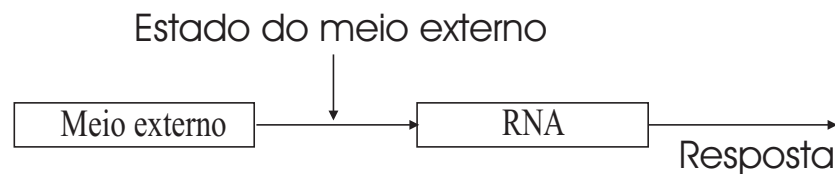


Figura 4.7. Aprendizado não-supervisionado.

4.5.3 Aprendizado por competição

O aprendizado por competição é um caso particular de aprendizado não-supervisionado. A ideia nesse caso é que para um dado padrão de entrada, deve-se fazer com que as unidades de saída disputem entre si para serem ativadas. Existe, portanto, uma competição entre as unidades de saída para decidir qual delas será a vencedora e, conseqüentemente, terá sua saída ativada e seus pesos atualizados no treinamento. Dessa forma, uma vez que uma saída tenha sido ativada a chance de vencer outra disputa aumenta e assim a unidade mais forte fica ainda mais forte. Esse tipo de situação também é conhecida como *winner-takes-all*.

4.5.4 Aprendizado por reforço

O aprendizado por reforço pode ser visto como um caso particular de aprendizado supervisionado, entretanto, embora o aprendizado supervisionado forneça as

entradas e saídas desejadas, no aprendizado por reforço a única informação de realimentação fornecida à rede é se uma determinada saída está ou não correta. No aprendizado supervisionado a medida de desempenho é baseada em um conjunto de respostas desejadas usando algum critério de erro conhecido. Já no aprendizado por reforço o desempenho é baseado em qualquer informação que possa ser fornecida ao sistema.

4.6 Redes perceptron e adaline

O perceptron é uma das formas mais simples de rede neural usada na classificação de padrões que sejam linearmente separáveis. Basicamente, o perceptron consiste em um único neurônio com pesos sinápticos ajustáveis. Uma rede perceptron (com um único neurônio) pode classificar apenas dois padrões distintos (que sejam linearmente separáveis). O algoritmo de aprendizado do perceptron pode ser visto na Equação (4.3):

$$w_i(t+1) = w_i(t) + \eta e x_i(t) , \quad (4.3)$$

sendo:

$$e = d - y . \quad (4.4)$$

O termo d representa a saída esperada da rede, enquanto y é a saída obtida pela rede. Portanto, e significa a diferença entre a saída esperada e a obtida, $w_i(t)$ é o valor de peso correspondente à entrada i no tempo t , η é o valor da taxa de aprendizado, e $x_i(t)$ representa o valor da entrada i no tempo t . Através dessa regra, todos os pesos devem ser corrigidos até que ocorra a convergência da rede, e esta apresente os resultados desejados a todas as entradas.

Paralelamente ao modelo perceptron, o modelo adaline (*adaptive linear neuron*), foi proposto por Widrow e Hoff [29]. A grande diferença entre os algoritmos de treinamento dos modelos perceptron e adaline consiste que no modelo adaline a correção dos pesos é baseada na saída linear (antes da função de ativação). Esse algoritmo de aprendizado (conhecido por regra delta) deu origem ao famoso algoritmo *backpropagation*, muito utilizado no treinamento de redes neurais multicamadas.

4.7 Redes MLP (*Multi Layer Perceptron*)

Conforme visto anteriormente, as redes de uma só camada resolvem apenas problemas linearmente separáveis, ou funções de limiar. Para entradas booleanas, as funções de limiar correspondem a um pequeno subconjunto do total de funções booleanas possíveis. A solução de problemas não linearmente separáveis passa pelo

uso de redes com uma ou mais camadas intermediárias ou escondidas. Redes com apenas uma camada intermediária podem implementar qualquer função contínua.

Segundo Braga [16], Redes MLP apresentam um poder computacional muito maior do que aquele apresentado pelas redes sem camadas intermediárias. Teoricamente, redes com duas camadas intermediárias podem implementar qualquer função, seja ela linearmente separável ou não. A precisão obtida e a implementação da função objetivo dependem do número de nodos utilizados nas camadas intermediárias.

Um dos principais aspectos relacionados ao projeto de redes MLPs diz respeito à função de ativação utilizada. Diversas funções de ativação têm sido propostas para redes multicamadas. Estas funções são não-lineares e diferenciáveis, e isso se deve ao fato de que, ao se utilizar uma função linear, apesar da utilização de uma ou várias camadas intermediárias, ainda haverá a restrição da linearidade, que é exatamente o problema a ser resolvido.

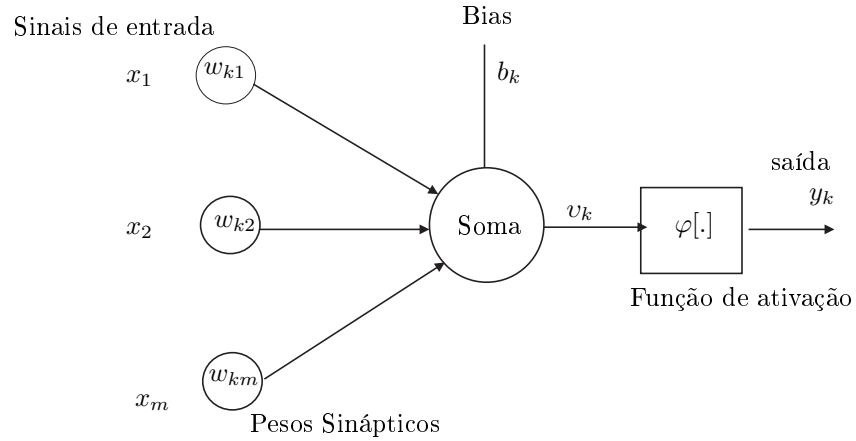
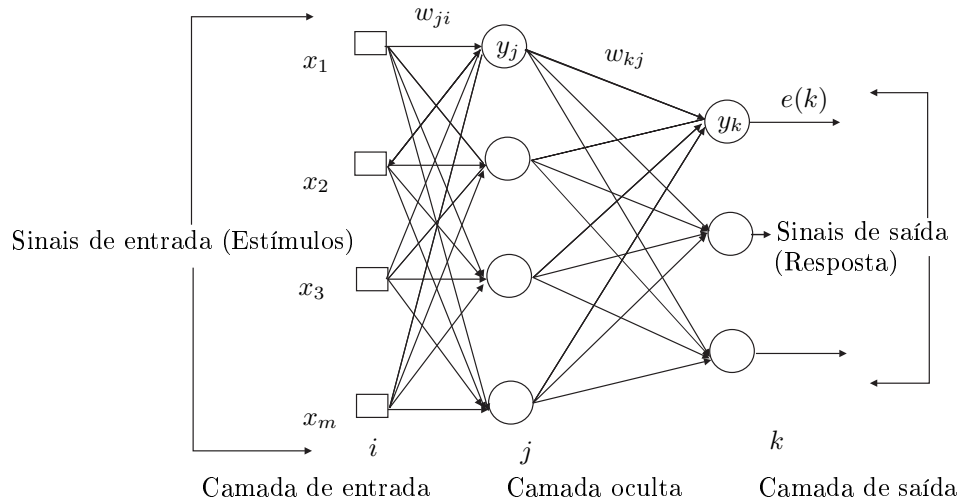
4.7.1 Arquitetura do perceptron de múltiplas camadas

Utilizando como estrutura o neurônio artificial como na Figura 4.8, forma-se a base para projetos de redes neurais artificiais, onde se podem observar três elementos básicos:

1. entrada, representado pelo produto interno de x_m e $w_{\kappa m}$. Assim, um sinal x_j , na entrada da sinapse j , conectado ao neurônio k , é multiplicado pelo peso sináptico $w_{\kappa j}$. O primeiro índice se refere ao neurônio em questão e o segundo se refere ao terminal de entrada ao qual o peso se refere;
2. um somador, que soma o produto interno dos sinais de entrada com os respectivos pesos dos neurônios;
3. uma função de ativação, ou de restrição, que restringe a amplitude da saída y_κ do neurônio.

O modelo neural da Figura 4.8 [30], possui também um limiar de ativação (*bias*), aplicado externamente, representado por b_κ . O *bias* b_κ tem o efeito de aumentar ou diminuir a entrada da função de ativação, dependendo se ele é positivo ou negativo, respectivamente.

Utilizando como estrutura básica da rede o neurônio da Figura 4.8, tipicamente a rede possui uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. Na Figura 4.9 [30] apresenta-se a arquitetura de uma rede perceptron de múltiplas camadas com uma camada de entrada, uma oculta e uma de saída. Nota-se que um neurônio em qualquer camada da rede está conectado a todos os neurônios da camada anterior.

**Figura 4.8.** Neurônio artificial.**Figura 4.9.** Arquitetura perceptron de múltiplas camadas com uma camada oculta.

4.7.2 Treinamento de redes multicamadas

Existem atualmente vários algoritmos para treinar redes MLP, que são geralmente do tipo supervisionado [16]. De acordo com os parâmetros que eles atualizam, os algoritmos para treinamento de redes do tipo MLP podem ser classificados em estáticos e dinâmicos. Enquanto os algoritmos estáticos não alteram a estrutura da rede, variando apenas os valores de seus pesos, os algoritmos dinâmicos podem tanto reduzir quanto aumentar o tamanho da rede (número de camadas, número de nodos nas camadas intermediárias e número de conexões). Quando o aprendizado estático é utilizado, a mesma regra de aprendizado é empregada para redes do tipo MLP com diferentes tamanhos e formatos. É interessante observar que topologias diferentes podem resolver o mesmo problema.

4.7.3 Algoritmo de retropropagação

O algoritmo de retropropagação de erro possui duas fases distintas:

- para frente ou propagação (*forward*). As características de uma amostra são introduzidas na camada de entrada, propaga-se para frente (neurônio por neurônio) através da rede e emerge na camada de saída da rede;
- para trás ou retropropagação (*backward*). O erros que se originam nos neurônios da camada de saída, se propagam para trás (camada por camada) por meio da rede, atualizando os seus pesos.

Na Figura 4.10, apresentam-se os dois tipos de fluxos identificados nesta rede. Para suavizar a carga matemática envolvida na derivação do algoritmo de retropropagação, primeiro apresenta-se um resumo das notações utilizadas na derivação.

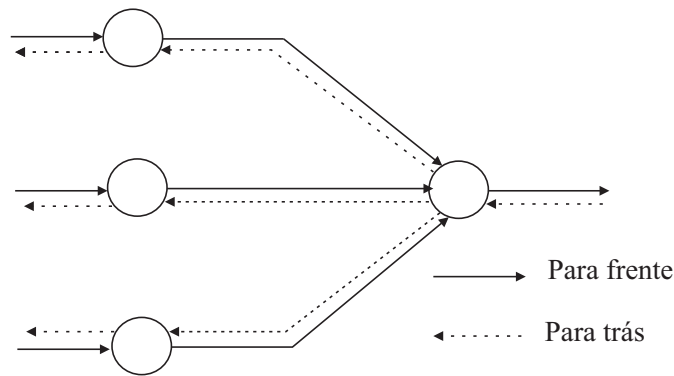


Figura 4.10. Ilustração das direções de dois fluxos básicos em um perceptron de múltiplas camadas: Propagação para frente e retropropagação.

- Os índices i , j e k apresentados na Figura 4.9 referem-se às camadas diferentes na rede; propagando através da rede da esquerda para a direita, a camada i se encontra na camada de entrada, a camada j como camada oculta e a camada k na camada de saída;
- N refere-se à quantidade de amostras do conjunto de treinamento;
- na iteração n , o n -ésimo padrão de treinamento (exemplo) é apresentado à rede;
- o símbolo $\epsilon(n)$ se refere à soma dos erros quadráticos na iteração n . A média de $\epsilon(n)$ sobre todos os valores de n (i.e., o conjunto inteiro de treinamento) produz o erro médio EQM ;
- o símbolo $e_{\kappa}(n)$ se refere ao erro na camada de saída do neurônio k , para a iteração n ;
- o símbolo $d_{\kappa}(n)$ se refere à resposta desejada para o neurônio k e é usado para calcular $e_{\kappa}(n)$;

- o símbolo $y_j(n)$ se refere à função de ativação que aparece na saída do neurônio j , na iteração n ;
- o símbolo $y_k(n)$ se refere à função de ativação que aparece na saída do neurônio k , na iteração n ;
- o símbolo $w_{ji}(n)$ representa o peso sináptico conectando a saída do neurônio i à entrada do neurônio j na iteração n . A correção aplicada a este peso na iteração n é representada por $\Delta w_{ji}(n)$;
- o i -ésimo elemento do vetor (padrão) de entrada é representado por $x_i(n)$;
- o parâmetro da taxa de aprendizagem é representado por η ;
- o símbolo m_1 representa o tamanho da camada de entrada, m_2 o tamanho da camada oculta e m_3 tamanho da camada de saída.

Na metodologia de Kevin [31] são descritos os seguintes passos para treinar uma rede neural, usando o algoritmo retropropagação.

1. Definir a arquitetura da rede. O tamanho da camada de entrada depende de quantas características o objeto possui, mais a unidade de limiar de ativação b_k . Na camada oculta, definida empiricamente, deve-se ter em conta que utilizando unidades demais pode-se levar a rede a memorizar os exemplos. Este problema é chamado de *overfitting*; de outro modo, com pequeno número pode ocorrer *underfitting*, no qual a rede não consegue convergir durante o treinamento. A camada de saída é a quantidade de classes nas quais o objeto pode ser classificado.
2. Inicializar os pesos w aleatoriamente entre $[-1, 1]$.
3. Inicializar as ativações das unidades que introduzem os limiares nas unidades da camada posterior: $x_0 = 1,0$ e $y_0 = 1,0$. Em alguns casos pode ser -1.
4. Entrar com um exemplo de treino $x_i(n)$ (no modo sequencial escolhido aleatoriamente do conjunto de treinamento).
5. Propagar a ativação das unidades de entrada, que são os próprios dados de entrada, para as unidades da camada oculta, usando a função de ativação sigmoidal.

$$y_j = \frac{1}{1 + \exp\left(-\sum_{i=1}^{m_1} w_{ji}x_i\right)}, \text{ para } j = 1, 2, \dots, m_2 \quad (4.5)$$

6. Propagar as ativações das unidades da camada oculta, para as unidades da camada de saída.

$$y_\kappa = \frac{1}{1 + \exp\left(-\sum_{j=1}^{m_2} w_{\kappa j} x_j\right)}, \text{ para } \kappa = 1, 2, \dots, m_3 \quad (4.6)$$

7. Calcular o erro $e_\kappa(n)$ na camada de saída do neurônio κ .

$$e_\kappa(n) = d_\kappa(n) - y_\kappa(n) \quad (4.7)$$

8. Calcular os deltas das unidades da camada de saída, denotada por δ_κ .

$$\delta_\kappa = y_\kappa(1 - y_\kappa)e_\kappa \quad \text{para } \kappa = 1, \dots, m_3 \quad (4.8)$$

9. Calcular os deltas das unidades da camada oculta, denotados por δ_j .

$$\delta_j = y_j(1 - y_j) \sum_{\kappa=1}^{m_3} \delta_\kappa w_{\kappa j} \quad \text{para } j = 1, \dots, m_2 \quad (4.9)$$

10. Calcular a correção para os pesos entre a camada oculta e a camada de saída:

$$\Delta w_{\kappa j} = \eta \delta_\kappa y_j \quad \text{para } \kappa = 0, \dots, m_2 \quad \text{e} \quad j = 1, \dots, m_3 \quad (4.10)$$

11. Calcular a correção para os pesos entre a camada de entrada e a camada oculta:

$$\Delta w_{ji} = \eta \delta_j x_i \quad \text{para } j = 0, 1, \dots, m_1 \quad \text{e} \quad i = 1, 2, \dots, m_2 \quad (4.11)$$

12. Ajustar os pesos entre a camada oculta e a camada de saída.

$$w_{\kappa j}(n+1) = w_{\kappa j}(n) + \Delta w_{\kappa j} \quad (4.12)$$

13. Ajustar os pesos entre a camada de entrada e a camada oculta.

$$w_{ji}(n+1) = w_{ji}(n) + \Delta w_{ji} \quad (4.13)$$

14. Ir para o passo 4 e repetir as operações.

Quando todos os padrões de entrada do conjunto de treinamento tiverem sido apresentados à rede, então uma época (*epoch*) foi completada. Deve-se repetir os passos de 4 a 14 para quantas épocas forem desejadas, ou se atingir um erro médio desejado.

Na próxima seção serão descritos os dois tipos de treinamento: sequencial e por lote.

4.8 Modo de treinamento seqüencial e por lote

Como mencionado anteriormente, uma apresentação completa do conjunto de treinamento é denominado época. Para um dado conjunto de treinamento segundo [30], a aprendizagem por retropropagação pode então proceder de uma entre duas formas básicas:

1. modo seqüencial. O modo seqüencial da aprendizagem por retropropagação é também chamado de modo *on-line*. Neste modo de operação a atualização dos pesos é realizada após a apresentação de cada exemplo de treinamento; este é o modo de apresentação para o qual se aplica a derivação do algoritmo de retropropagação apresentado. Para ser específico é preciso considerar uma época constituindo-se de N vetores exemplos de treinamento organizados na seguinte ordem $(x_1, d_1), \dots, (x_N, d_N)$. O primeiro par de exemplo (x_1, d_1) da época é apresentado à rede; e a seqüência de cálculos, para frente e para trás, descrita anteriormente, é realizada resultando em ajustes dos pesos sinápticos e níveis de *bias* da rede. Então, o segundo par de exemplo (x_2, d_2) da época é apresentado e a seqüência de computações para frente e para trás é repetida, resultando em novos ajustes dos pesos sinápticos e níveis dos limiares de ativação. Esse processo continua, até que o último par de exemplo (x_N, d_N) da época, seja considerado;
2. modo por lote. No modo por lote da aprendizagem por retropropagação o ajuste dos pesos é realizado após a apresentação de todos os exemplos de treinamento que constituem uma época.

Do ponto de vista operacional *on-line*, o modo seqüencial de treinamento é preferível em relação ao modo por lote, porque requer menos armazenamento local para cada conexão sináptica. Além disso, dado que os parâmetros são apresentados à rede de uma forma aleatória, o ajuste de pesos de padrão em padrão torna a busca no espaço de pesos de natureza aleatória. Isso pode diminuir as possibilidades do algoritmo de retropropagação ficar preso em um mínimo local [30].

O modo de treinamento por lote fornece uma estimativa precisa do vetor gradiente. A convergência para um mínimo local é assim garantida sob condições simples. A composição do modo por lote também o torna mais fácil de ser paralelizado do que o modo seqüencial.

Quando o conjunto de dados de treinamento contém várias cópias exatas dos mesmos padrões, verifica-se que, diferentemente do modo por lote, o modo seqüencial é capaz de tirar vantagens de sua redundância porque os exemplos são apresentados um de cada vez. Isso ocorre particularmente quando o conjunto de dados é grande e altamente redundante [30].

4.8.1 Vetor gradiente

Admitindo que a função $f:D \subset R^2 \rightarrow R$ é diferenciável em $P_0 \in D$, define-se vetor gradiente da função f no ponto $P_0 = (x, y)$ e representa-se por $\nabla f(P_0)$ ou $\text{grad}f(P_0)$ ao vetor cujas componentes são as derivadas parciais de f no ponto P_0 , ou seja,

$$\nabla f(P_0) = \frac{\partial f(P_0)}{\partial x}, \frac{\partial f(P_0)}{\partial y}. \quad (4.14)$$

Uma importante propriedade do gradiente é que ele aponta na direção de maior crescimento de f em (x, y) . O vetor oposto do gradiente (gradiente descendente):

$$-\nabla f(x, y) = \frac{-\partial f}{\partial x}(x, y), \frac{-\partial f}{\partial y}(x, y), \quad (4.15)$$

aponta para a direção em que a função decresce mais rapidamente.

4.8.2 Gradiente descendente

De acordo com a metodologia estabelecida por Haykin [30], o algoritmo de retro-propagação utiliza o gradiente descendente para minimizar a função de custo ϵ que é a soma dos erros ao quadrado instantânea da rede, pois no modo seqüencial isso ocorre a cada apresentação do vetor x . A função de custo ϵ , é escrita como:

$$\epsilon(n) = \frac{1}{2} \sum_{\kappa=1}^{m_3} e_{\kappa}^2(n), \quad (4.16)$$

sendo m_3 o número de neurônios da camada de saída da rede e e_{κ} o erro do neurônio κ , na iteração n .

O cálculo dos δ para a correção dos pesos depende da posição da camada. Considerando que o neurônio analisado encontra-se na camada de saída κ , sendo estimulado por um conjunto de ativações produzido por uma camada de neurônios à sua esquerda, a saída $v(n)$ do neurônio κ é:

$$v_{\kappa}(n) = \sum_{j=1}^{m_2} w_{\kappa j}(n) y_j(n). \quad (4.17)$$

A função de ativação $y_{\kappa}(n)$ que se encontra na saída do neurônio κ na interação n é:

$$y_{\kappa}(n) = \varphi_{\kappa}(v_{\kappa}(n)). \quad (4.18)$$

Diferenciando a Equação (4.18) em função a $v_{\kappa}(n)$ 4.17 obtém-se:

$$\frac{\partial y_{\kappa}(n)}{\partial v_{\kappa}(n)} = \varphi'_{\kappa}(v_{\kappa}(n)). \quad (4.19)$$

O algoritmo de retropropagação aplica uma correção $\Delta w_{\kappa j}$ ao peso sináptico $w_{\kappa j}$, que é proporcional à derivada parcial $\partial \epsilon(n)/\partial w_{\kappa j}(n)$. Tem-se então que a correção ao peso sináptico pode ser expressa por:

$$\Delta w_{\kappa j}(n) = \eta \delta_{\kappa}(n) y_{\kappa}(n) , \quad (4.20)$$

sendo o gradiente local $\delta_{\kappa}(n)$ definido por:

$$\delta_{\kappa}(n) = e_{\kappa}(n) \varphi'_{\kappa}(v_{\kappa}(n)) . \quad (4.21)$$

Segundo a Equação (4.21), o gradiente local $\delta_{\kappa}(n)$ para o neurônio de saída κ é igual ao produto do erro $e_{\kappa}(n)$ correspondente para aquele neurônio pela derivada $\varphi'_{\kappa}(v_{\kappa}(n))$ da função de ativação associada. Sendo que o neurônio κ está localizado na camada de saída da rede, ele é alimentado com uma resposta desejada particular. Pode-se utilizar a Equação (4.7) para calcular o erro $e_{\kappa}(n)$ associado a este neurônio. Tendo-se determinado $e_{\kappa}(n)$ calcula-se diretamente o gradiente local $\delta_{\kappa}(n)$, utilizando a Equação (4.21). De acordo com Haykin [30] o gradiente local aponta para as modificações necessárias nos pesos sinápticos.

Quando o neurônio j está localizado na camada oculta da rede, não existe uma resposta desejada especificada para aquele neurônio. Assim, o erro para um neurônio oculto deve ser determinado de forma recursiva, na forma dos erros de saída de todos os neurônios, aos quais o neurônio oculto está diretamente conectado. De acordo com a Equação (4.21), o gradiente local $\delta_j(n)$ para o neurônio oculto j é representado na Equação 4.22:

$$\delta_j(n) = -\frac{\partial \epsilon(n)}{\partial y_j(n)} \varphi'_j(v_j(n)) , \quad (4.22)$$

sendo que para calcular a derivada parcial $\partial \epsilon(n)/\partial y_j(n)$, tem-se a Equação 4.23:

$$\epsilon(n) = \frac{1}{2} \sum_{\kappa=1}^{m_3} e_{\kappa}^2(n) , \quad (4.23)$$

diferenciando a Equação 4.23 em relação a ativação $y_j(n)$, obtém-se:

$$\frac{\partial \epsilon(n)}{\partial y_j(n)} = \sum_{\kappa=1}^{m_3} e_{\kappa} \frac{\partial e_{\kappa}(n)}{\partial y_j(n)} . \quad (4.24)$$

A seguir, utiliza-se a regra da cadeia para a derivada parcial $\partial e_{\kappa}(n)/\partial y_j(n)$ e rescrevendo a Equação 4.24 obtém-se a Equação 4.25:

$$\frac{\partial \epsilon(n)}{\partial y_j(n)} = \sum_{\kappa=1}^{m_3} e_{\kappa}(n) \frac{\partial e_{\kappa}(n)}{\partial v_{\kappa}(n)} \frac{\partial v_{\kappa}(n)}{\partial y_j(n)} . \quad (4.25)$$

A partir deste ponto, obtém-se a derivada parcial desejada:

$$\frac{\partial \epsilon(n)}{\partial y_j(n)} = - \sum_{\kappa} \delta_{\kappa}(n) w_{\kappa j}(n) . \quad (4.26)$$

Utilizando a definição de gradiente local δ_κ abordada anteriormente chega-se ao gradiente local $\delta_j(n)$ apresentado segundo a Equação 4.27.

$$\delta_j(n) = \varphi'_j(v_j(n)) \sum_{\kappa=1}^{m_3} \delta_\kappa(n) w_{\kappa j}(n). \quad (4.27)$$

4.8.3 Modo por lote

Considerando a arquitetura da Figura 4.9 completamente conectada tem-se que o Erro Quadrático Médio (EQM), a ser minimizado, é dado por:

$$EQM = \frac{1}{2N} \sum_{n=1}^N \sum_{\kappa=1}^{m_3} e_\kappa^2(n). \quad (4.28)$$

Na Equação (4.28), a somatória interna em relação a κ é realizada sobre todos os neurônios da camada de saída da rede, enquanto que o somatório externo em relação a n é realizado sobre todo o conjunto de treinamento da época considerada.

Para calcular a fórmula do ajuste de pesos no modo por lote é preciso calcular a derivada da Equação (4.28) em relação aos pesos da primeira camada w_{ji} e em relação aos pesos da segunda camada w_{kj} , pois $EQM = f(w_{ji}, w_{kj})$ para todos i, j e κ . Através disso tem-se que o ajuste ao peso sináptico $w_{\kappa j}$, depois de uma época, é dado por:

$$\Delta w_{\kappa j} = -\frac{\eta}{N} \sum_{n=1}^N \delta_\kappa(n) y_j(n). \quad (4.29)$$

Observa-se que é preciso calcular $e_\kappa(n)$ e δ_κ para cada n e depois fazer a média do produto dessas quantidades multiplicada pela taxa de aprendizado η .

4.9 Taxa de aprendizado e constante de momento

A taxa de aprendizado é uma variável utilizada para medir a rapidez com que o vetor de pesos será atualizado. Segundo Haykin [30] o algoritmo de retropropagação fornece uma aproximação para a trajetória no espaço de pesos calculada pelo método da descida mais íngreme. Quanto menor for a taxa de aprendizado η , menor será a variação dos pesos sinápticos da rede, de uma iteração para a outra, e mais suave será a trajetória no espaço de pesos. Essa situação fornece uma melhoria ao algoritmo sobre o custo de uma taxa de aprendizado lenta. Caso a taxa de aprendizado η for muito grande, para acelerar a taxa de aprendizagem, a rede pode-se tornar instável devido às grandes modificações nos pesos sinápticos.

A constante de momento é uma variável que pode ser adicionada ao algoritmo de retropropagação, na equação de ajuste de pesos, com o objetivo de aumentar a velocidade de aprendizado, reduzindo possíveis situações de instabilidade. A regra

delta da Equação (4.20) modificada pela inclusão de uma constante de momento é apresentada como:

$$\Delta w_{ji}(n) = \alpha \Delta w_{ji}(n-1) + \eta \delta_j(n) y_i(n), \quad (4.30)$$

sendo que α é usualmente um número positivo chamado de constante de momento.

A inclusão da constante de momento no algoritmo de retropropagação representa uma pequena alteração na atualização dos pesos; no entanto, alguns resultados favoráveis ao comportamento de aprendizagem do algoritmo podem ser observados. A constante de momento pode favorecer o algoritmo evitando que o processo de aprendizagem termine em um mínimo local na superfície de erro.

4.10 Função de ativação

Um dos principais aspectos para projetos de redes perceptron multicamadas diz respeito à função de ativação. Como visto anteriormente, essas funções devem ser não lineares e diferenciáveis. Um exemplo de uma função de ativação não-linear, continuamente diferenciável, normalmente utilizada nos perceptrons de múltiplas camadas é a sigmóide logística não-linear. Para o cálculo do δ para cada neurônio do perceptron de múltiplas camadas utiliza-se a derivada da função de ativação $\varphi(\cdot)$ associado àquele neurônio. Para existir a derivação, a função $\varphi(\cdot)$ deve ser contínua. Basicamente, a diferenciabilidade é a única exigência que a função deve satisfazer. Essa sigmóide na sua forma geral é definida por:

$$\varphi_j(v_j(n)) = \frac{1}{1 + \exp(-av_j(n))}, \quad a > 0 \text{ e } -\infty < v_j(n) < +\infty \quad (4.31)$$

sendo $v_j(n)$ a função do neurônio j . De acordo com essa não-linearidade, a saída se encontra dentro do intervalo $0 \leq y_j \leq 1$. Diferenciando a Equação (4.31) em relação a $v_j(n)$, tem-se:

$$\varphi_j'(v_j(n)) = \frac{a \exp(-av_j(n))}{[1 + \exp(-av_j(n))]^2}. \quad (4.32)$$

Com $y_j = \varphi_j(v_j)$, pode-se eliminar o termo exponencial $\exp(-av_j(n))$ da Equação (4.32) e assim expressar a derivada $\varphi_j'(v_j(n))$ como:

$$\varphi_j'(v_j(n)) = a y_j(n) [1 - y_j(n)]. \quad (4.33)$$

4.11 Critério de parada

A convergência de um algoritmo de retropropagação, em geral, não pode ser demonstrada e não existem regras bem definidas para finalizar sua operação. Um

dos critérios de parada utilizados tem como objetivo calcular a função de custo ou medida de erro $EQM(w)$, que é estacionária no ponto de valor ótimo do vetor de peso. Infelizmente, este critério pode resultar em um encerramento prematuro do processo de aprendizagem.

Também é possível comparar a função de custo ou medida de erro $EQM(w)$, com uma constante ϕ : se $EQM(w) < \phi$ pode-se encerrar o treino. Quando a soma dos erros quadráticos $SSE = \sum_{k=1}^m (d_k - y_k)^2$ aumentar por n vezes seguidas o treinamento também é encerrado.

Outro critério de parada é o *cross-validation*, o qual, após cada iteração do aprendizado, a rede é testada pelo seu desempenho de generalização. O processo de aprendizagem é encerrado quando o desempenho de generalização for adequado, ou quando ficar aparente que o desempenho de generalização atingiu o máximo.

Outro critério de parada ocorre quando o algoritmo treina até um número pré-determinado de épocas. O ideal é usar combinações entre os métodos.

4.12 Representação da saída e regra de decisão

Utilizando (x, d) como par de treinamento, sendo x o vetor de entrada, d a resposta desejada de x , supondo que as amostras possuam apenas duas classes A e B e que a representação da saída utilizada no treinamento for normalizada para valores entre 0 e 1 (que eventualmente correspondem aos limites superior e inferior das saídas da rede, quando a função logística é usada), pode-se representar que d_{n1} pertence à classe A e d_{n2} à classe B

$$d_{n1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

$$d_{n2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

De acordo com Haykin [30] uma regra de decisão de saída apropriada é a regra Bayesiana, onde tomando $y_1(n)$ e $y_2(n)$ como saídas produzidas pela rede ao se apresentar um vetor aleatório x , esse vetor pertencerá à classe A quando:

$$y_1(n) > y_2(n), \quad (4.34)$$

caso contrário pertencerá à classe B .

É interessante notar que os valores-alvo binários $[0, 1]$, correspondentes à função logística da Equação (4.31), são alterados por uma pequena quantidade ϵ , como uma medida prática para evitar a saturação dos pesos sinápticos (devido à precisão numérica finita) durante o treinamento da rede. Como resultado desta alteração, os

valores-alvo são agora não-binários. Abaixo, um exemplo em vez de 0 e 1 será 0,2 e 0,8.

$$d_{n1} = \begin{bmatrix} 0,8 \\ 0,2 \end{bmatrix}$$
$$d_{n2} = \begin{bmatrix} 0,2 \\ 0,8 \end{bmatrix}$$

4.13 Divisão do conjunto de dados

Não existe uma regra clara para divisão do conjunto de dados, o que se pode fazer é utilizar heurísticas onde o próprio projetista da rede deverá decidir a melhor forma para a divisão.

Segundo Lutz [32], podem ser destinadas 50% das amostras para os dados de treinamento, 25% para dados de validação, que é utilizado como condição de parada e para evitar *overfitting* e 25% para dados de testes, simulando futuras coletas e são utilizados durante o treinamento da rede. Na divisão do conjunto de dados, deve-se selecionar para cada classe de forma aleatória as amostras colocando-as em suas respectivas matrizes.

4.14 Validação cruzada

A validação cruzada ou *cross-validation*, tem como princípio ajudar na parada do treinamento. Primeiramente, divide o conjunto de dados aleatoriamente em conjunto de treinamento e conjunto de teste. O conjunto de treinamento é dividido adicionalmente em dois subconjuntos separados:

- subconjunto de estimação, utilizado para o ajuste dos parâmetro da rede;
- subconjunto de validação, utilizado para testar ou validar o modelo.

De forma usual, a rede é treinada utilizando o subconjunto de estimação, exceto por uma pequena modificação: o treinamento é interrompido de forma periódica após um determinado número de épocas, e a rede é testada com o subconjunto de validação, isto é, após um período de treinamento, os pesos sinápticos e os níveis do limiar de ativação (*bias*) do perceptron de múltiplas camadas são todos fixos, e a rede opera no seu modo direto, para frente. O erro de validação é então medido para cada exemplo do subconjunto de validação. Quando a fase de validação é completada, a estimação (treinamento) é reiniciada para um novo período, e o processo é repetido.

Na Figura 4.11 apresenta-se as características conceituais de duas curvas de aprendizagem: uma referente às medidas sobre o subconjunto de estimação, e a outra

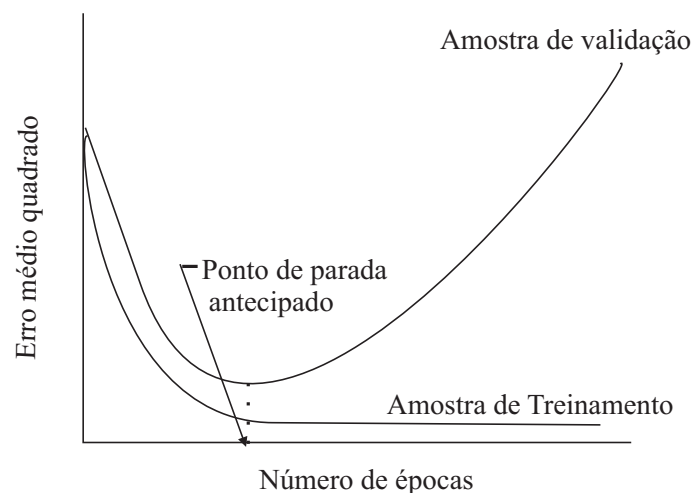


Figura 4.11. Ilustração da regra de parada antecipada baseada na validação cruzada.

referente ao subconjunto de validação. Segundo Haykin [30], a curva de aprendizagem de estimação decresce para um número crescente de épocas. Diferentemente, a curva de aprendizagem de validação decresce para um mínimo e então começa a crescer conforme o treinamento continua. Ao observar a curva de aprendizagem de estimação pode parecer que pode-se melhorar o desempenho indo além do ponto mínimo da curva de aprendizagem de validação. Entretanto, na realidade, o que a rede aprende após esse ponto é essencialmente o ruído contido nos dados de treinamento. Essa heurística sugere que o ponto mínimo na curva de aprendizagem de validação seja usado como critério para encerrar a sessão de treinamento. Esse procedimento é referido como método de treinamento com parada antecipada [30].

4.15 Generalização

A rede neural utiliza o conjunto de treinamento para adaptar os pesos e o limiar de ativação (*bias*) da rede, de forma que a rede neural seja capaz de generalizar. Uma rede generaliza bem quando o mapeamento de entrada-saída computado pela rede for correto para dados de teste que não foram utilizados para o treinamento da rede. Assume-se que os dados de teste são retirados da mesma população usada para gerar os dados de treinamento.

Uma rede neural pode ter uma generalização pobre, quando há neurônios ocultos mais do que necessário, resultando em memorização. De outro modo, com pequeno número pode ocorrer *underfitting*, onde a rede não consegue convergir durante o treinamento. Deve-se escolher entre o limite de *overfitting* e *underfitting* [16].

Na Figura 4.12 apresenta-se um exemplo de generalização que pode acontecer em uma rede neural. O mapeamento não-linear de entrada-saída representado pela curva é computado pela rede como resultado da aprendizagem dos pontos rotulados

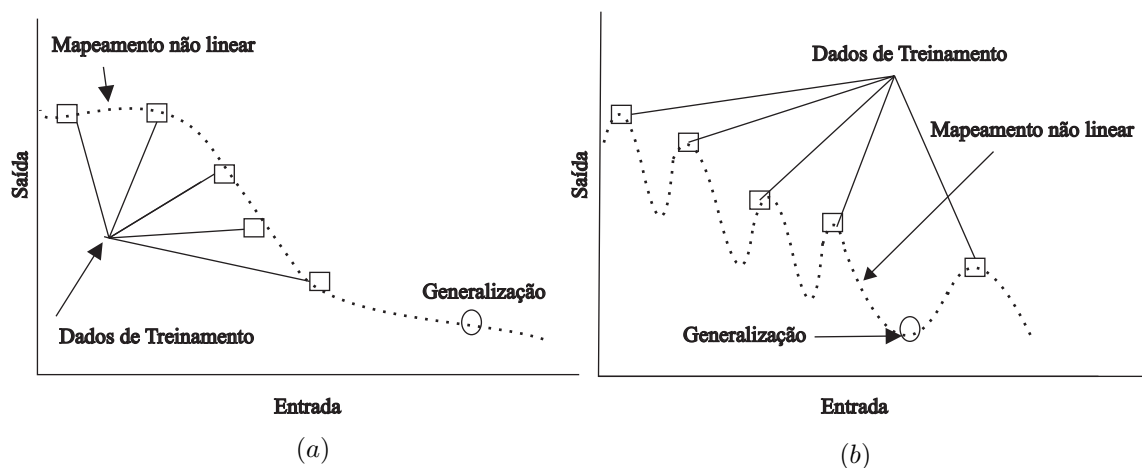


Figura 4.12. *a* -Dados ajustados adequadamente; *b* -Dados ajustados em excesso.

como dados de treinamento e o ponto marcado como generalização é visto como resultado da interpolação realizada pela rede. Uma memorização em uma rede neural implica que o mapeamento de entrada-saída da rede não é suave [30]. Na Figura 4.12a observa-se que ocorreu uma boa generalização, ou seja, os dados foram ajustados de forma adequada. Já na Figura 4.12b observa-se que ocorreu uma generalização pobre, devido a memorização, os dados foram ajustados em excesso.

4.15.1 Avaliação da eficiência de um algoritmo de aprendizagem

Um algoritmo de aprendizagem é considerado bom se produz hipóteses que permitem prever satisfatoriamente as classes a que pertencem exemplos que não tenham sido vistos anteriormente. A seguir está descrita uma metodologia para avaliar a qualidade de previsão depois que as amostras foram apresentadas, conforme descritos por Russel e Norvig [33].

Uma previsão é considerada boa se é confirmada; por isso, para avaliar a qualidade de uma hipótese, verificam-se suas previsões em relação à classificação correta, uma vez que é conhecida. Isso se realizará em um conjunto de exemplos conhecidos como conjunto de teste. Se para realizar treinamento se utilizam todos os exemplos de que se dispõe, então deve-se que buscar outros para efetuar provas ou testes. Devido a isso, frequentemente é preferível adotar a seguinte metodologia:

1. coletar uma grande quantidade de exemplos;
2. dividi-la em dois conjunto distintos: o conjunto de treino e o conjunto de teste;
3. aplicar o algoritmo de aprendizagem com o conjunto de treino para produzir uma hipótese H ;

4. calcular a porcentagem de exemplos do conjunto de teste corretamente classificado com H ;
5. repetir os passos 1 a 4 em conjuntos de treino de tamanhos para diferentes e seus exemplos escolhidos aleatoriamente.

Nota-se que, conforme aumenta o conjunto de treinamento (padrões significativos), também pode-se aumentar a precisão da classificação. Esse comportamento da curva de aprendizagem é um bom indício de que, nos dados de fato, subentende-se um padrão no qual o algoritmo de aprendizagem está incorporando.

4.16 Heurística para melhorar o desempenho

Projetos de uma rede neural utilizando o algoritmo de retropropagação, usualmente são mais reconhecidos como uma arte do que uma ciência, isso em virtude dos inúmeros fatores envolvidos no projeto. Esses fatores podem ser absolvidos a partir da experiência particular de cada um. No entanto, existem métodos que visam melhorar o desempenho do algoritmo de retropropagação, como descrito [30]:

1. tornar aleatório o conjunto de treinamento. Nas tarefas de classificação de padrões usando a aprendizagem por retropropagação seqüencial, uma técnica simples bastante utilizada é tornar aleatória (i.e., embaralhar) a ordem em que os exemplos são apresentados ao perceptron de múltiplas camadas de uma época para a seguinte. Aleatoriar pode garantir que os exemplos apresentados à rede em uma época raramente pertençam à mesma classe;
2. valores-alvo. Trabalhar com os valores alvo (resposta desejada) dentro do intervalo da função de ativação sigmóidal. Mais especificamente, a resposta desejada d_j para o neurônio j na camada de saída do perceptron de múltiplas camadas deve ser deslocada por uma quantidade ϵ afastada do valor limite da função de ativação sigmóide, dependendo se o valor limite é positivo ou negativo. Caso contrário, o algoritmo de retropropagação tende a levar os parâmetros livres da rede para o infinito e dessa forma reduzir a velocidade do processo de treinamento, levando os neurônios ocultos à saturação;
3. normalizar as entradas. Cada variável de entrada deve ser pré-processada de modo que o seu valor médio, calculado sobre todo o conjunto de treinamento seja próximo de zero, ou seja pequeno comparado com o desvio padrão.
4. variar o número de neurônios na camada oculta;
5. a taxa de aprendizado é escolhida experimentalmente para cada problema;
6. a inicialização afeta muito o resultado final. Reinicializações podem resolver o problema.

4.17 Técnica de poda de rede

Técnicas de poda ou *pruning* envolvem a eliminação de pesos e neurônios irrelevantes para a função executada pela rede. No método baseado na avaliação da sensibilidade da saída, é retirado um neurônio da rede e verificada a variação do erro de saída. Caso a retirada da unidade de processamento não cause grande variação no erro de saída, conclui-se que a rede é pouco sensível a este, e que ele pode ser retirado sem perda da capacidade da rede modelar os dados [16].

4.18 Virtudes e limitações

De acordo com Haykin [30], há vantagens e desvantagens do algoritmo de retropropagação de erro.

4.18.1 Vantagens

O algoritmo de retropropagação baseia-se em cálculos locais para descobrir as capacidades de processamento de informações por redes neurais. Esta forma de restrição computacional é referida como a restrição de localidade, ou seja, a computação realizada pelo neurônio é influenciada apenas por aqueles neurônios que estão em contato físico [30]. O uso de computação local pela unidade de processamento, permite uma degradação suave no desempenho devido a erros dos sensores e, portanto, fornece base para um projeto de rede tolerante a falhas. A computação local é de natureza maciçamente paralela.

A complexidade computacional do algoritmo *backpropagation* é polinomial em relação ao ajuste dos parâmetros (pesos) que devem ser atualizados de uma iteração para a seguinte. Nesse sentido o algoritmo *backpropagation* é computacionalmente eficiente.

4.18.2 Desvantagens

O algoritmo de retropropagação de erro é basicamente uma técnica de tarefas que podem ser complexas e lentas. A existência de mínimos locais (vales) pode gerar uma falsa solução ótima, ou aumentar o tempo de treinamento para sair do vale.

Os problemas de treinamentos de redes neurais em grande escala são considerados tão difíceis que não existe uma estratégia de aprendizagem que seja única, podendo ser necessárias outras abordagens, como por exemplo o uso de pré-processamento, algoritmos genéticos, lógica nebulosa, etc.

4.19 Comentários

O capítulo apresentou conceitos sobre a rede neural utilizada nesse trabalho, redes MLP ou multi-camadas. Foram explanadas arquiteturas, treinamento, o algoritmo de treinamento (*Backpropagation*), função de ativação, e regras de tomada de decisão. Algumas peculiaridades necessárias para um melhor treinamento de redes MLP também foram discutidas e apresentadas. Todo este estudo realizado sobre teorias de redes MLP foi de grande importância para o entendimento e desenvolvimento de uma solução para o problema proposto neste trabalho.

Capítulo 5

Métodos experimentais e resultados

Neste capítulo são apresentados os procedimentos e os resultados obtidos para caracterização de amostras de leite por métodos convencionais, métodos por ultra-som e os resultados para as redes neurais treinadas.

A metodologia apresentada no capítulo de técnicas convencionais é aplicada para se obter os resultados para calibração das amostras de leite UHT. Essa calibração tem como objetivo obter o teor de gordura, quantidade de água adicionada e densidade em amostras calibradas para uma posterior comparação com os parâmetros obtidos pelas técnicas de ultra-som.

O leite UHT foi utilizado, pois é um líquido que apresenta características homogêneas e possui especificações vindas de fábrica. Ele pode ser caracterizado com maior facilidade por se tratar de um líquido que possui distribuição e formato de partículas regulares. Outro fato importante, a ser observado, é que ao se utilizar os métodos apresentados, pode-se inspecionar o leite UHT após seu processamento, o que é realizado em laticínios, como controle de qualidade da produção.

O treinamento das redes neurais foi realizado a partir da escolha de algumas configurações de redes neurais. Essas configurações foram determinadas variando os dados de quantidade de nodos ocultos, taxa de aprendizado, constante de momento quantidade de iterações. Após o treinamento das diversas configurações utilizadas, foi observado o desempenho dessas redes, e a partir dessa etapa, foram escolhidas as melhores redes projetadas. Numa fase final, as redes neurais foram submetidas a testes com amostras desconhecidas pela rede.

5.1 Preparação e calibração de amostras

Os experimentos foram realizados com amostras de leite UHT (*Ultra high temperature*) integral e desnatado do mesmo fabricante e do mesmo lote, que foram utilizados com o objetivo de variar a proporção de gordura, realizando misturas entre eles. A Tabela 5.1 apresenta as informações nutricionais para o leite UHT integral comercial para cada 200 ml. A composição do leite UHT integral é prati-

Tabela 5.1. Informação nutricional: Porção de 200 ml para o leite UHT integral.

Componente	Massa (%)
Gorduras totais	3,3
Gorduras saturadas	1,6
Proteínas	3,2
Carboidratos	4,6
Sódio	0,06
Colesterol	0,001
Cálcio	0,13
Ferro	0,00018

camente igual à do leite desnatado, diferindo basicamente na proporção de gordura. Isto foi examinado através de experimentos convencionais com o objetivo de calibrar as amostras e relacioná-las com os parâmetros acústicos.

Os experimentos com o objetivo de calibrar as amostras foram realizados a uma temperatura média de $25 \pm 1^\circ\text{C}$. As amostras com proporções de gordura utilizadas foram obtidas realizando dosagens entre os leites (integral e desnatado). Estas dosagens foram realizadas utilizando uma balança eletrônica, que possui uma sensibilidade e reprodutibilidade de 0,01g.

Na intenção de observar variações na composição do leite UHT integral por adição de água, foram preparadas várias amostras com diferentes quantidades de água adicionada ao leite UHT integral e essas amostras também são analisadas.

Os experimentos foram realizados de acordo com as metodologias disponíveis e estudadas no capítulo 2. Esses experimentos foram realizados no laboratório de Biotecnologia da FE-UNESP-Campus de Ilha Solteira, de responsabilidade da professora Jacira dos Santos Isepon. Foram determinados o teor de gordura, a quantidade de água adicionada e densidade, que foram utilizados como dados de calibração para os experimentos por ultra-som.

Utilizou-se o método de Gerber para a obtenção do teor de gordura das misturas entre o leite UHT integral e desnatado. A Tabela 5.2, mostra os resultados obtidos, pelo método de Gerber em temperatura ambiente, para concentração de gordura das amostras em volume e o desvio médio para cada conjunto de amostras. A resolução do método de Gerber é de 0,1%. Da mesma forma, a Tabela 5.3 apresenta os dados de concentração de gordura, para amostras de leite UHT com adição de água. Observa-se que a concentração de gordura decai proporcionalmente com a adição de água. Assim obtêm-se amostras calibradas que são relacionadas com os parâmetros acústicos, velocidade, atenuação e densidade. Dessa relação são levantadas curvas utilizadas para apresentar as dependências entre os parâmetros.

Para determinação dos resultados obtidos para o teor de gordura, quantidade de água adicionada e densidade foram realizadas seis medidas para cada amostra, sendo

Tabela 5.2. Teor de gordura (ϕ_v) obtido pelo método de Gerber, em temperatura ambiente (26°C) para diluições de leite UHT em porcentagens diferentes.

Leite UHT	ϕ_v [%]
Integral	$3,0 \pm 0,01$
Integral + desnatado	$2,9 \pm 0,05$
Integral + desnatado	$2,0 \pm 0,05$
Integral + desnatado	$1,6 \pm 0,05$
Integral + desnatado	$1,0 \pm 0,05$
Integral + desnatado	$0,6 \pm 0,01$
Desnatado	$0,5 \pm 0,05$

Tabela 5.3. Teor de gordura (ϕ_v) obtido pelo método de Gerber, em temperatura ambiente (26°C) para o leite UHT integral com diferentes porcentagens de água adicionada.

Água adicionada ao leite integral [%]	ϕ_v [%]
60	$1,1 \pm 0,02$
50	$1,4 \pm 0,01$
20	$2,3 \pm 0,01$
10	$2,6 \pm 0,03$
5	$2,7 \pm 0,08$
2	$2,8 \pm 0,04$
1	$2,9 \pm 0,03$

a partir destes obtidos a média e os desvios apresentados nas tabelas 5.2, 5.3, 5.4 e 5.5. Para cada amostra foram realizadas duas medidas da mesma para confirmar o resultado.

Na Figura 5.1 apresenta-se a variação da concentração de gordura em função da adição de água. Observa-se que a concentração de gordura decai proporcionalmente com a adição de água, formando uma curva aproximadamente linear. Observou-se que a quantidade de água no leite UHT não é alterada após a diluição de amostras com diferentes concentrações de gordura.

A densidade foi medida através do termolactodensímetro e picnômetro. A Tabela 5.4 apresenta os resultados obtidos pelos métodos. Os resultados mostram que esses métodos não possuem discriminação significativa para o leite com pequenas variações de gordura. Isso ocorre devido à relativa proximidade entre as densidades dos glóbulos de gordura (870 kg/m^3) e do soro do leite (1030 kg/m^3), e com as baixas concentrações, as variações são da ordem do desvio na medida da densidade.

A diminuição da densidade do leite com a adição de água foi comprovado ex-

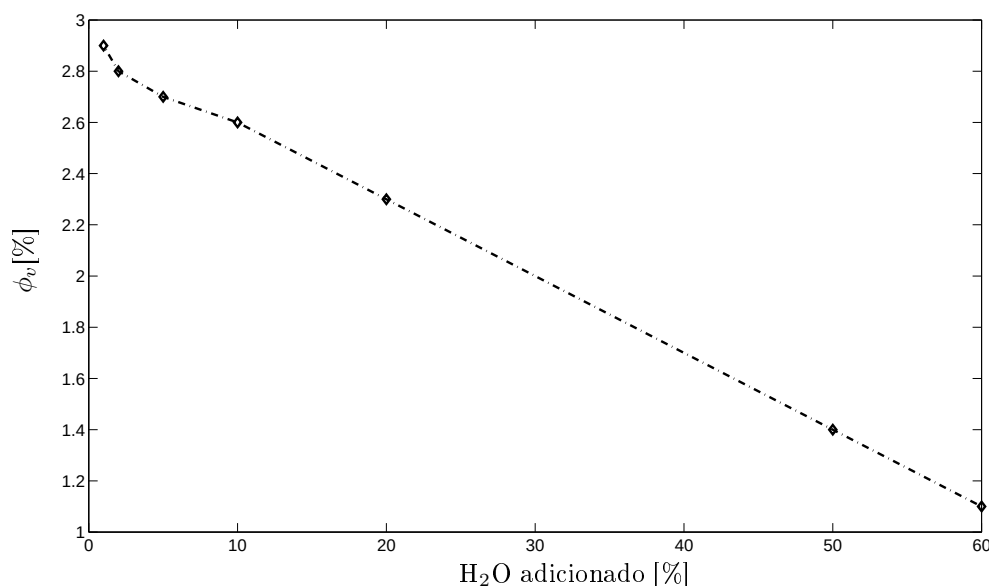


Figura 5.1. Concentração de gordura em função da quantidade de água adicionada.

Tabela 5.4. Densidades obtidas pelos métodos que utilizam o termolactodensímetro e picnômetro para diferentes concentrações de gordura à temperatura média de 26°C.

ϕ_v [%]	ρ_{pic} [kg/m ³]	ρ_{termo} [kg/m ³]
2,9	1027,3 ± 0,1	1031 ± 0,3
1,6	1028,5 ± 0,03	1031,8 ± 0,5
0,9	1029,2 ± 0,04	1032 ± 0,7
0,6	1028,5 ± 0,05	1032 ± 1,0
0,5	1028 ± 0,02	1033 ± 0,8

perimentalmente através dos métodos que utilizam o termolactodensímetro e o picnômetro. Esses resultados obtidos são mostrados na Tabela 5.5.

5.2 Descrição do equipamento de ultra-som

Nas Figuras 5.2 e 5.3 são apresentados o esquema dos equipamentos utilizados nos experimentos por ultra-som e a foto da montagem, respectivamente. Utiliza-se um pulsador/receptor de ultra-som (Panametrics 5077PR) que é responsável pela excitação do transdutor de frequência central de 5 MHz, gerando um pulso estreito, e pela amplificação do sinal proveniente do transdutor receptor. Os sinais do receptor são amplificados e filtrados no 5077PR e digitalizados por um osciloscópio digital Tektronix modelo TDS2022 de dois canais, com suporte para interfaces GPIB e RS232. A resolução do osciloscópio é de 8 bits, sua taxa de amostragem é de

Tabela 5.5. Densidades obtidas pelos métodos que utilizam o termolactodensímetro e picnômetro, para diferentes quantidades de água à temperatura média de 25°C .

Água adicionada [%]	$\rho_{pic}[\text{kg/m}^3]$	$\rho_{termo}[\text{kg/m}^3]$
60	$1008,6 \pm 0,4$	$1009,0 \pm 0,5$
50	$1011,8 \pm 0,4$	$1015,0 \pm 0,3$
20	$1022 \pm 1,0$	$1023,6 \pm 0,4$
10	$1025,1 \pm 0,4$	$1026,6 \pm 0,2$
5	$1026,6 \pm 0,4$	$1028,6 \pm 0,8$
2	$1027,4 \pm 0,4$	$1029 \pm 0,6$
1	$1027,7 \pm 0,6$	$1030 \pm 0,4$

2 Amostras/s, e a largura de banda é 200 MHz.

São utilizados sensores de temperatura, os quais são colocados em diversos pontos da célula de medição. Esses sensores são do tipo RTDs (*Resistance Temperature Detectors*), que baseiam-se na variação da resistência elétrica em função da temperatura. Os RTDs utilizados são de platina do tipo: PT-100 ($R_0 = 100 \Omega$) e PT-1000 ($R_0 = 1000 \Omega$). Utiliza-se uma placa de aquisição de temperatura National Instruments modelo NI4351. A resolução da placa é de 24 bits e esta funciona em conjunto com o dispositivo externo TBX-68T. O equipamento TBX-68T é um bloco de acesso aos terminais da placa de aquisição e possui 15 canais de aquisição, 2 fontes de corrente ($25 \mu\text{A}$ e 1 mA) e 8 saídas digitais. A placa de aquisição mede valores de tensão entre os terminais de seus canais, assim, utiliza-se fontes de corrente quando deseja-se medir resistências.

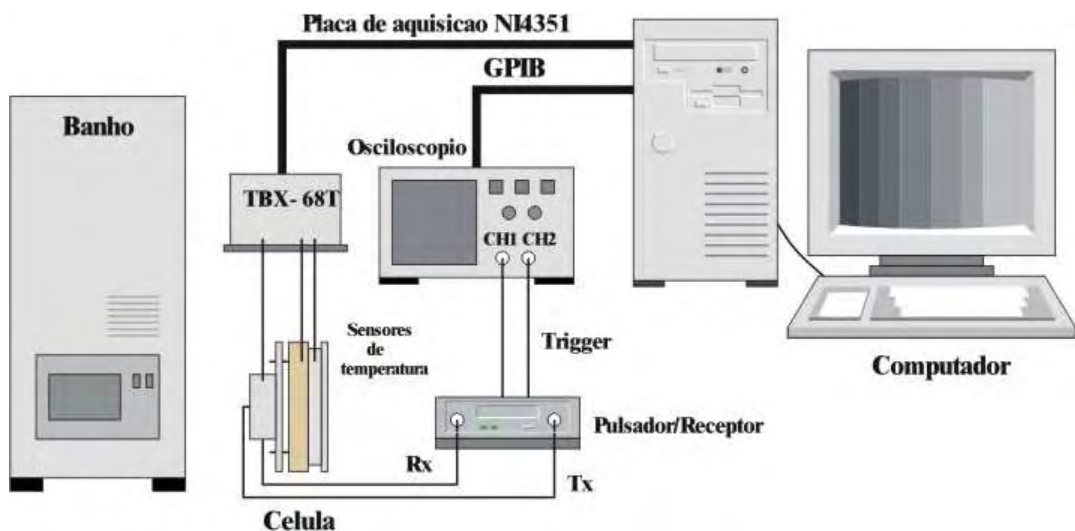


Figura 5.2. Esquema experimental.

A placa de aquisição e o osciloscópio são controlados por um programa desenvolvido em MatlabTM, onde também é realizado o processamento e exibição dos

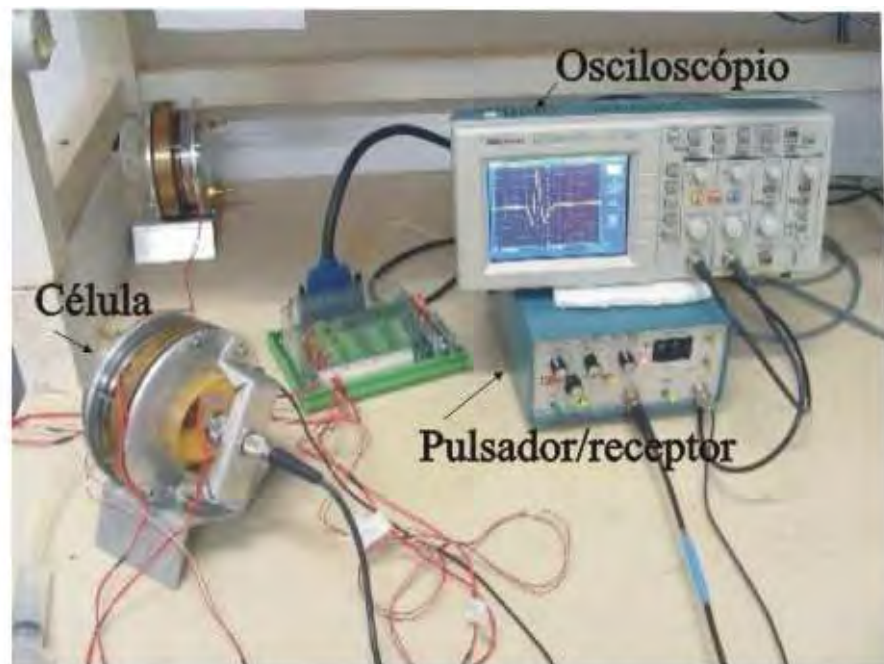


Figura 5.3. Foto do sistema de medição .

dados.

O banho ultratermostatizado (Marconi, modelo MA 184), foi utilizado para experimentos com temperatura controlada, cuja temperatura varia entre -10 e 100°C com erros de $\pm 0,12\%$. É possível trabalhar com rampas de temperatura para obter a relação dos parâmetros acústicos e do leite em função da variação de temperatura de controle.

5.3 Calibração da célula

Antes de iniciar a medição dos parâmetros acústicos das amostras, a célula deve passar pelo processo de calibração, cujo o objetivo é determinar o comprimento L_2 da câmara de amostra, e minimizar os erros devido a problemas de alinhamento entre as faces e ruídos elétricos. Utiliza-se a água destilada para a calibração da célula, pois é um líquido que possui características acústicas bem conhecidas em função da temperatura [34, 35]. Como a água destilada possui a velocidade de propagação c_2 conhecida em função da temperatura, mede-se a diferença de tempo entre os sinais, Δt_{12} e a partir desta medição calcula-se o comprimento da câmara de amostra L_2 . O valor de L_2 obtido é utilizado posteriormente para realizar as medições de velocidade de propagação do leite [13].

O comprimento da câmara de amostra L_2 é corrigido em função da variação de temperatura. Observa-se uma variação da câmara de amostra em torno de $0,2 \mu\text{m}/^{\circ}\text{C}$. Para corrigir o comprimento da câmara de amostra a uma determi-

nada temperatura T mede-se o comprimento de calibração L_{2cal} a uma temperatura de calibração T_{cal} e então utiliza-se a Equação 5.1 [36]:

$$L_2(T) = L_{2cal} + (T - T_{cal})\Delta_L. \quad (5.1)$$

Para comprovar que a célula foi calibrada corretamente faz-se a medição da densidade da água destilada, utilizando a Equação (3.13). O valor obtido para a densidade é comparado com valor tabelado para uma determinada temperatura de operação (999 Kg/m³ para 26 °C). Ao se obter um erro relativo na ordem de 1%, entre os valores tabelado e medido, a célula de medição é considerada calibrada.

5.4 Resultados: métodos de ultra-som

Os métodos para caracterização de amostras de leite por ultra-som utilizam a relação experimental entre os parâmetros acústicos e as características do leite. Assim são verificadas as dependências da velocidade de propagação da onda (c), do coeficiente de atenuação (α) e a densidade (ρ), das amostras de leite UHT, em função da temperatura, frequência, teor de gordura e quantidade de água adicionada.

Os experimentos por ultra-som foram realizados mediante extremos cuidados com a manipulação das amostras, manuseio do equipamento de instrumentação e controle de temperatura. O controle de temperatura se faz de grande importância, pois a temperatura exerce grande influência sobre os parâmetros acústicos. Assim, foram realizados experimentos com rampas de temperatura para investigar essa relação entre os parâmetros acústicos e a temperatura. É importante destacar o longo tempo de execução de experimentos com rampas de temperatura, que geralmente demandaram de 8 a 10 horas para cada amostra.

5.4.1 Variação do teor de gordura

A velocidade de propagação da onda (c) é relacionada com o teor de gordura do leite UHT, e através disto analisou-se a dependência entre esses dois parâmetros em função de uma variação de temperatura entre 15°C e 35°C. Na Figura 5.4 apresenta-se os resultados para essa relação. Observa-se que a velocidade (c) decai com o aumento da concentração de partículas de gordura no meio. Quanto maior a concentração de partículas de gordura menor será a velocidade de propagação. Pode-se observar que existem faixas de temperatura nas quais obtém-se uma maior discriminação dos valores de velocidade, em função do teor de gordura do leite. Na Figura 5.5 apresenta-se uma ampliação da Figura 5.4, onde observa-se uma maior sensibilidade na faixa de temperatura entre 25°C e 35°C. Para caracterizar as demais amostras de leite foram realizados experimentos na faixa de temperatura entre 25°C e 35°C.

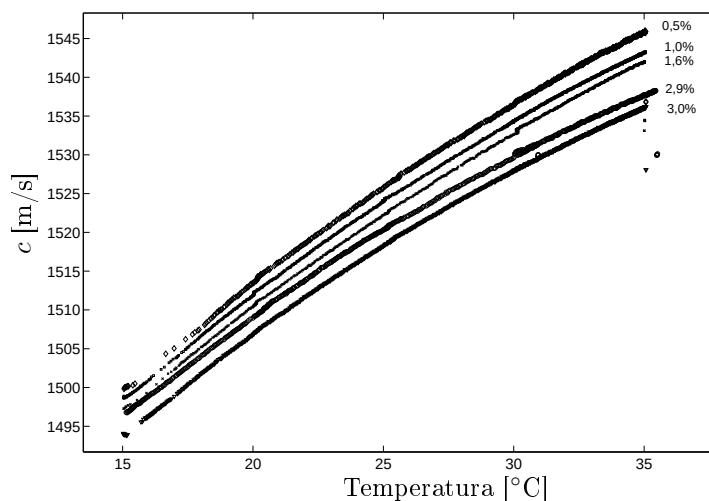


Figura 5.4. Velocidade de propagação da onda no leite UHT em função da temperatura para diferentes concentrações de gordura.

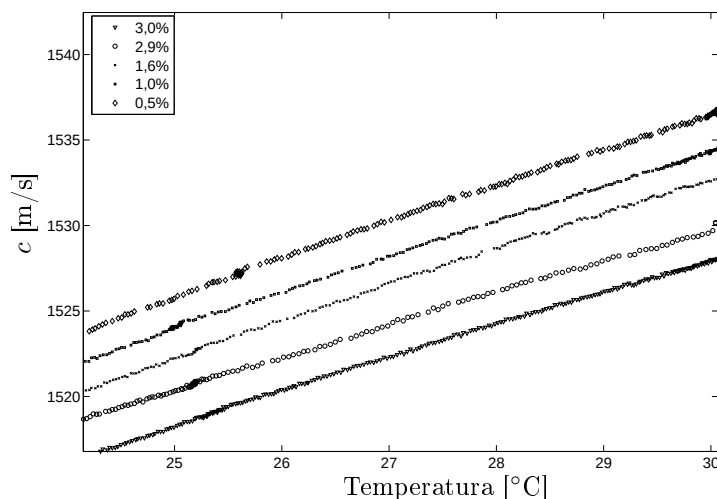


Figura 5.5. Velocidade de propagação da onda no leite UHT, ampliada, em função da temperatura para diferentes concentrações de gordura.

Na Figura 5.6 apresenta-se a velocidade de propagação da onda (c) em função do teor de gordura do leite UHT para uma temperatura de $26,5^{\circ}\text{C}$. Observa-se que a curva possui uma característica aproximadamente linear e a velocidade (c) decai com o aumento da concentração de partículas de gordura no meio. Obtém-se uma variação na velocidade, igual a $-3,0 \text{ (m/s)/ \% gordura}$, considerando que o leite UHT desnatado, contenha uma concentração de gordura de aproximadamente $0,5\%$ e o leite UHT integral $3,0\%$, como medido experimentalmente pelo método de Gerber. Os desvios nas medidas podem chegar a $0,09 \text{ m/s}$.

A atenuação da onda (α) também é outro parâmetro acústico que é utilizado da mesma forma que a velocidade de propagação da onda, ou seja, relaciona-se a aten-

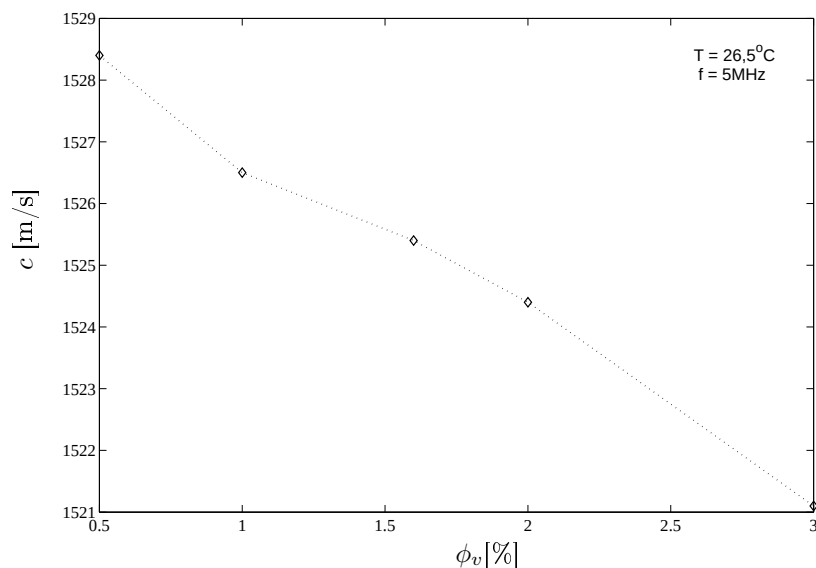


Figura 5.6. Velocidade de propagação da onda no leite UHT em função da concentração de gordura.

uação com o teor de gordura. A atenuação na amostra aumenta proporcionalmente com o aumento da concentração de partículas de gordura no meio. Observa-se que o leite UHT possui atenuações muito baixas, pois é um líquido de baixa concentração [10]. A atenuação foi normalizada pelo leite desnatado, ou seja, os valores indicam quanto a atenuação é superior à atenuação no leite desnatado. Na Figura 5.7 apresenta-se a variação da atenuação com o aumento de concentração de gordura no leite UHT. Obtém-se uma variação na atenuação, igual a 7,6 (Np/m)/% gordura, com desvios que podem chegar a 0,9 Np/m. Os resultados obtidos apresentam uma clara dependência entre o coeficiente de atenuação e o teor de gordura para amostras de leite UHT. A Figura 5.8 mostra que a atenuação possui um comportamento aproximadamente linear e torna-se mais significativa de acordo com o aumento da frequência, ou seja, têm-se uma maior sensibilidade para o coeficiente de atenuação em frequência maiores. Trabalhando com um transdutor com uma maior frequência resultados mais significativos podem ser obtidos. A Figura 5.8 também fortalece a dependência da atenuação em relação a concentração de partículas na amostra de leite, ou seja, quanto maior a concentração de gordura no leite maior será a atenuação sobre a onda de ultra-som.

Na Figura 5.9 apresenta-se o coeficiente de atenuação para o leite UHT para diferentes concentrações de gordura em função da temperatura. Observa-se que esse parâmetro apresentou pequenas variações na faixa de temperatura apresentada. Essas pequenas variações são resultados da baixa concentração das amostras e da proximidade dos valores. Uma pequena variação nos valores de atenuação representa um fator positivo para essa variável, pois ela depende menos da temperatura que a velocidade.

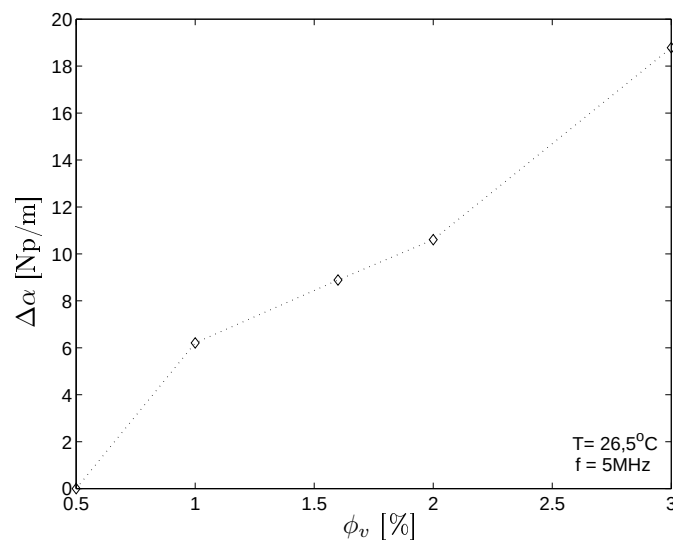


Figura 5.7. Coeficiente de atenuação da onda no leite UHT para frequência de 5 MHz em função do teor de gordura.

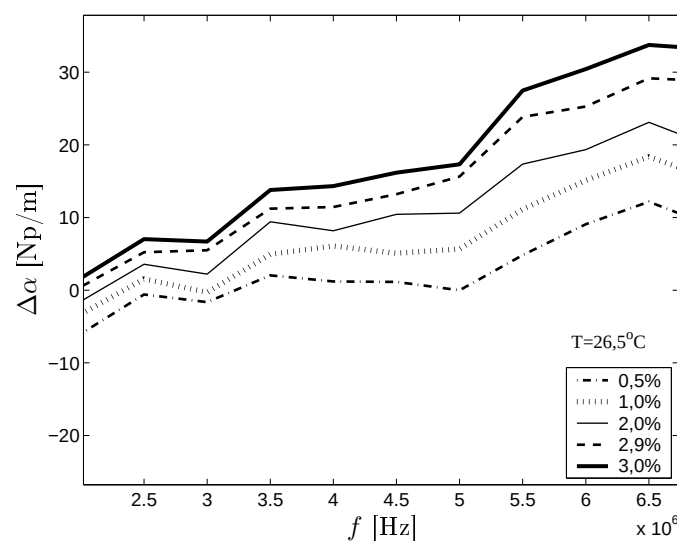


Figura 5.8. Coeficiente de atenuação da onda no leite UHT em função da frequência, para diferentes concentrações de gordura.

A densidade da amostra de leite UHT é calculada de forma direta de acordo com a Equação (3.13). Nota-se, pela Figura 5.10, que a densidade não apresentou variações significativas em função do teor de gordura. Isso porque, segundo Higuti [13], as densidades dos glóbulos de gordura (870 Kg/m^3) e do soro do leite (1030 Kg/m^3), são próximas e com baixas concentrações, as variações são da ordem do erro na medida da densidade. A Figura 5.11 mostra que a densidade também apresenta pequenas variações em função da temperatura, isso devido as características de formação do leite UHT descritas anteriormente.

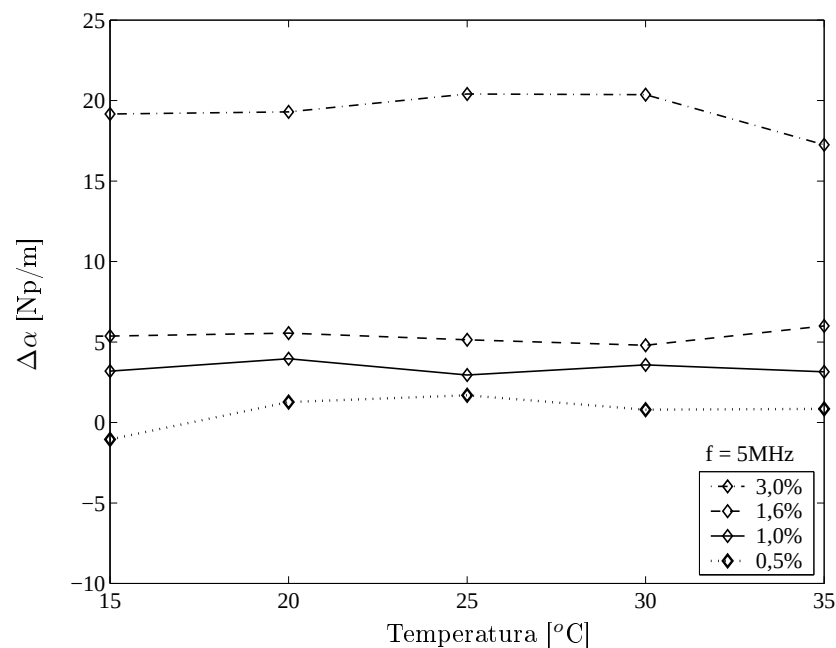


Figura 5.9. Coeficiente de atenuação da onda no leite UHT para frequência de 5 MHz em função da temperatura.

5.4.2 Variação da quantidade de água adicionada

A relação entre a velocidade de propagação da onda e a quantidade de água adicionada ao leite foi investigada. Ao conhecer a quantidade de água no leite, obtida através do método que usa o crioscópio, pode-se relacionar a velocidade de propagação com a porcentagem de água no leite UHT. A velocidade de propagação da onda (c) e a quantidade de água adicionada ao leite UHT (água adicionada [%]) em função de uma variação de temperatura entre 25°C e 35°C é apresentada na Figura 5.12. Pode-se observar que existem faixas de temperatura nas quais obtêm-

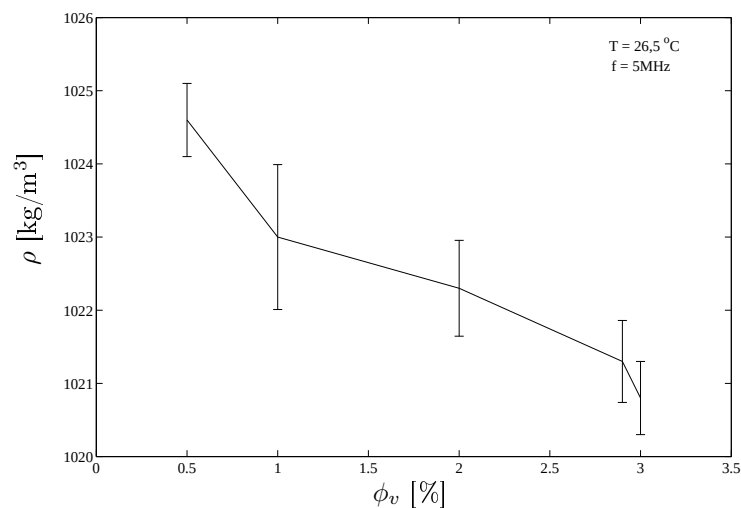


Figura 5.10. Densidade do leite UHT em função da concentração de gordura.

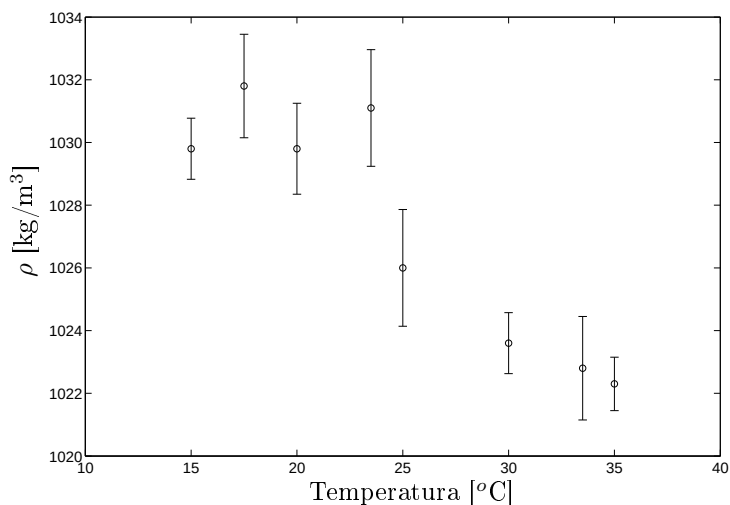


Figura 5.11. Densidade do leite UHT com 3,0% de teor de gordura (ϕ_v) em função da Temperatura.

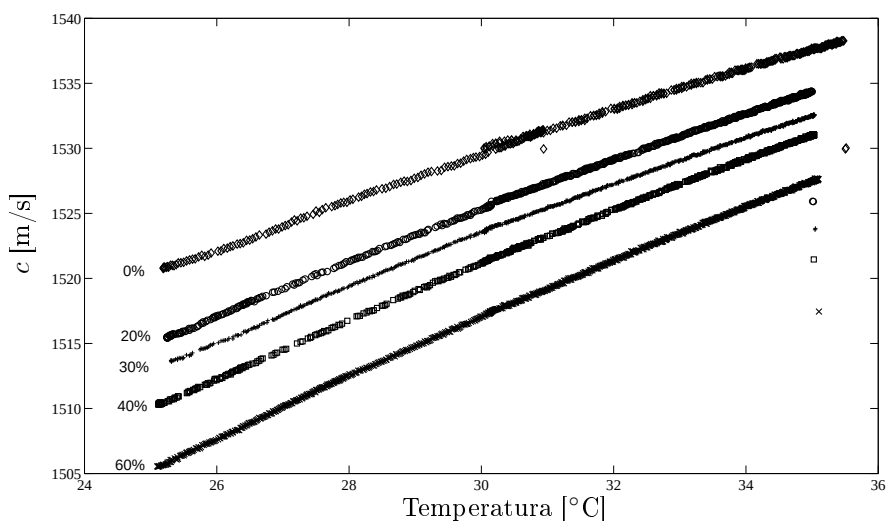


Figura 5.12. Velocidade de propagação da onda no leite UHT com diferentes quantidades de água adicionada em função da temperatura.

se uma maior discriminação dos valores de velocidade, em função da quantidade de água adicionada assim como para a análise anterior para o teor de gordura do leite. Como apresentado anteriormente a temperatura apresenta-se como uma variável importante no processo de discriminação do leite, pois modifica as propriedades acústicas do mesmo.

Na Figura 5.13 apresenta-se a relação entre velocidade de propagação da onda e a quantidade de água adicionada para uma temperatura de 26°C. Uma variação na velocidade igual a - 0,24 (m/s)/ % água foi obtida. Pode-se observar que ao adicionar água destilada ao leite UHT integral, ocorre uma diminuição na velocidade de propagação, isso devido a água destilada possuir uma velocidade menor que a do

leite integral.

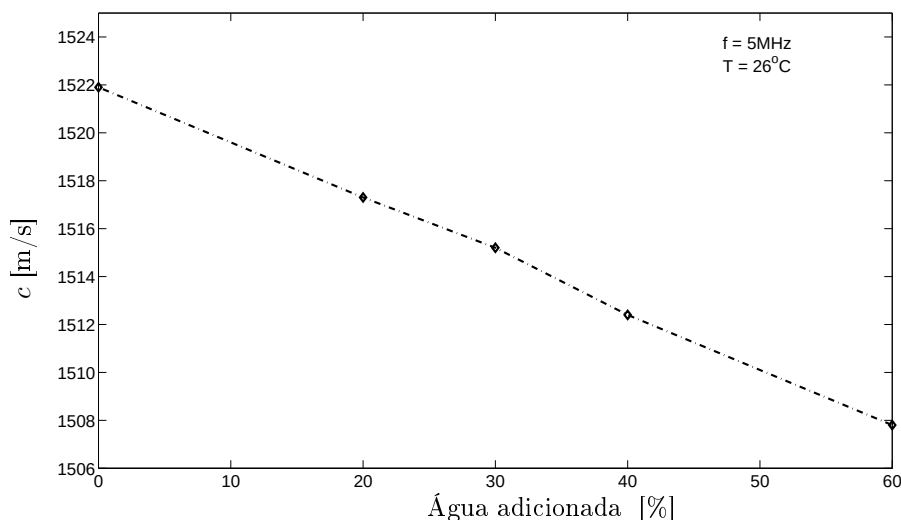


Figura 5.13. Velocidade de propagação da onda no leite UHT em função de água (H_2O) adicionada.

A atenuação da onda também foi utilizada para discriminar a quantidade de água no leite UHT. Relaciona-se a atenuação com a porcentagem de água no leite, e curvas que representam essa relação podem ser levantadas. A atenuação foi normalizada pelo leite com maior quantidade de água adicionada (60%), ou seja, os valores indicam quanto a atenuação é superior à atenuação no leite com 60% de água adicionada. Através dessas curvas observa-se que a atenuação sofre uma diminuição com o aumento na proporção de água destilada adicionada, isso devido à diminuição na concentração de partículas de gordura no leite. Na Figura 5.14 são apresentados esses resultados. Pode-se obter uma sensibilidade de $0,15 \text{ (Np/m)}/\%$ de água. Na Figura 5.15 apresenta-se os resultados da atenuação para várias frequências em função da quantidade de água adicionada ao leite. Nessa Figura (5.15) observa-se que pode-se obter resultados mais significativos trabalhando em uma faixa de frequência maior.

A densidade pode ser utilizada com maior poder de discriminação para amostras de leite UHT integral, com adição de água. Assim, pode-se usar a densidade para caracterizar esse processo de adulteração de leite por adição de água. Na Figura 5.16 apresenta-se os resultados obtidos para adição de água destilada no leite. Nota-se que ao adicionar água ao leite sua densidade diminui, isso pelo fato da densidade da água destilada ser menor que a densidade do leite integral.

5.5 Comparação com métodos convencionais

Nas Figuras 5.10, 5.11 e 5.16 são apresentados os resultados obtidos para variação da densidade do leite, em função da variação de gordura e em função da quantidade

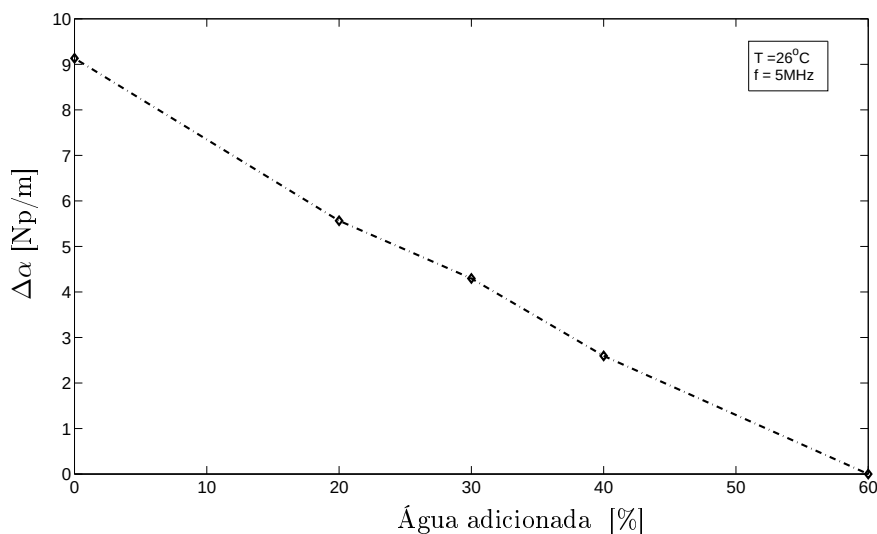


Figura 5.14. Coeficiente de atenuação da onda no leite UHT em função da quantidade de água adicionada .

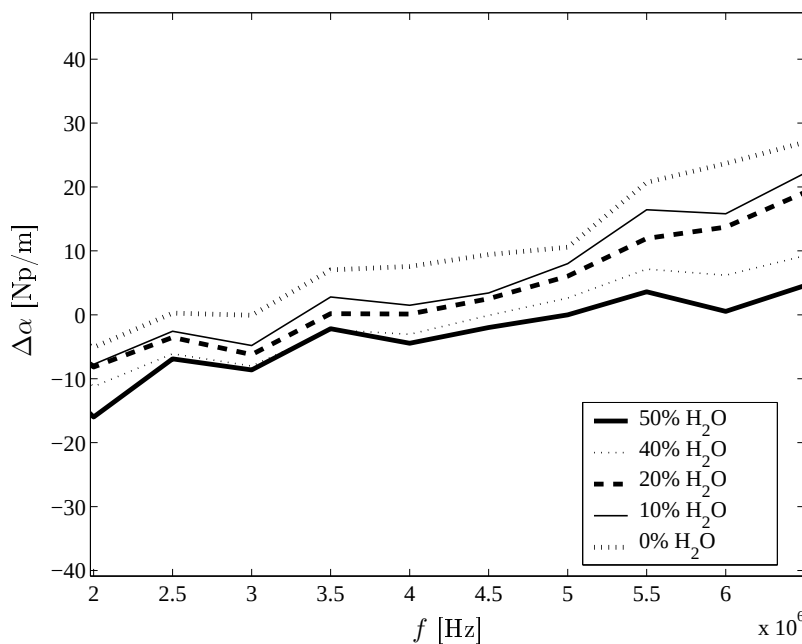


Figura 5.15. Coeficiente de atenuação da onda no leite UHT para diversas frequências em função da quantidade de água adicionada .

de água adicionada ao leite, respectivamente. Esses dados são provenientes dos experimentos realizados por ultra-som. No primeiro momento pode-se comparar esses resultados com resultados obtidos por métodos convencionais, que utilizam o termolactodensímetro e o picnômetro. Para os experimentos por métodos convencionais e por ultra-som foram realizados os mesmos procedimentos para a preparação de amostras. Foram utilizados amostras provenientes do mesmo fabricante, com as mesmas características, lote e faixa de temperatura.

Os resultados obtidos para densidade do leite em amostras com adição de água,

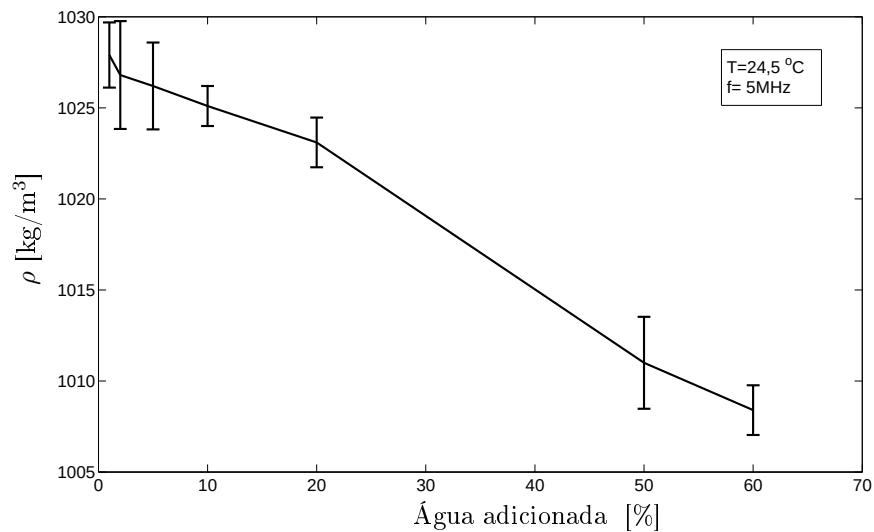


Figura 5.16. Densidade para o leite UHT integral em função de água adicionada.

Tabela 5.6. Densidades obtidas pelos métodos de ultra-som, Termolactodensímetro e picnômetro para variação de teor de água adicionada (a.a.)

a.a. [%]	$\rho_{cel}[\text{kg/m}^3]$	$\rho_{termo}[\text{kg/m}^3]$	$\rho_{pic}[\text{kg/m}^3]$
60	$1008,4 \pm 0,5$	$1009,0 \pm 0,3$	$1008,6 \pm 0,4$
50	$1011,0 \pm 0,6$	$1015,0 \pm 0,3$	$1011,8 \pm 0,4$
20	$1023,1 \pm 1,2$	$1023,6 \pm 0,4$	$1022,9 \pm 1,1$
10	$1025,1 \pm 1,5$	$1026,6 \pm 0,2$	$1025,1 \pm 0,4$
5	$1026,2 \pm 0,1$	$1028,6 \pm 0,6$	$1026,6 \pm 0,4$
2	$1026,8 \pm 0,8$	1029 ± 1	$1027,4 \pm 0,4$
1	$1027,9 \pm 0,2$	$1030 \pm 0,4$	$1027,7 \pm 0,6$

pelos métodos convencionais e por ultra-som, são apresentados e comparados na Tabela 5.6 e na Figura 5.17. Observa-se que os resultados obtidos pela célula de medição e picnômetro possuem uma proximidade significativa, considerando os desvios padrões das medidas. Por outro lado, nota-se uma maior discrepância entre os resultados obtidos por ultra-som e termolactodensímetro.

5.6 Métodos e resultados: redes neurais

Os dados obtidos pela célula de medição, velocidade e coeficiente de atenuação acústica em função da temperatura são utilizados por uma rede neural do tipo MLP, para classificar os teores de gordura do leite UHT entre 0,5 e 3,0%, com resolução de 0,1%. Utiliza-se uma rede MLP com 3 nodos na camada de entrada, representando os dados de temperatura, velocidade e atenuação da onda na frequência de 5 MHz

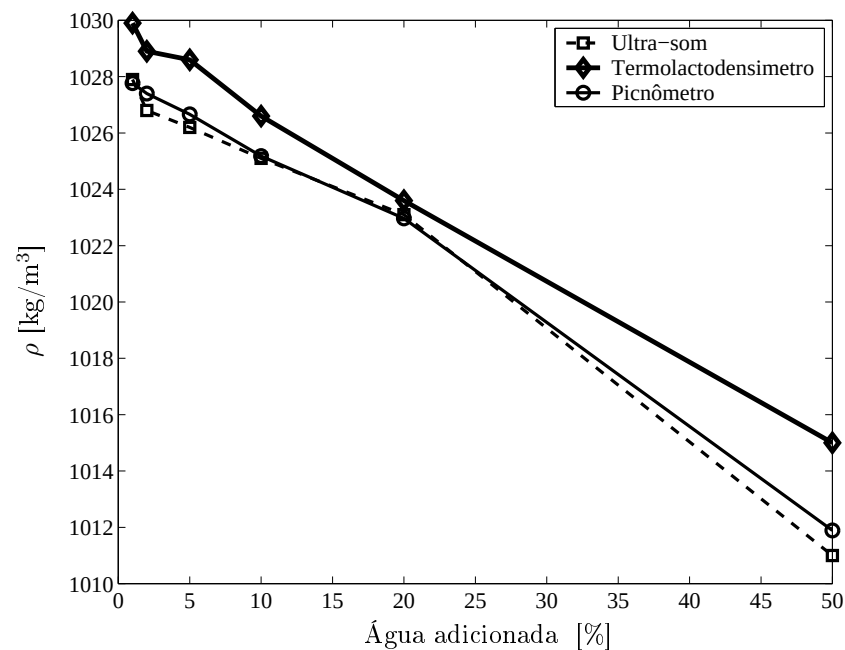


Figura 5.17. Densidade do leite UHT integral, para os métodos convencionais (Termolactodensímetro e picnômetro) e por ultra-som ($f=5\text{MHz}$) em função de água adicionada.

pela amostra de leite, e vinte e seis nodos na camada de saída e uma camada oculta variando o número de nodos em 26, 40 e 52. Os nodos da camada de saída representam os teores de gordura presentes no leite. Sendo que cada nodo representa um tipo de leite com um determinado teor de gordura. Na figura 5.18 apresenta-se a configuração de saída para a rede projetada para detecção da quantidade de gordura no leite. Utiliza-se também uma outra configuração com 5 nodos na camada de entrada, representando os dados de temperatura, velocidade e atenuação da onda na faixa de 4 MHz, 5 MHz e 6,5 MHz.

A densidade foi um parâmetro utilizado com o objetivo de caracterizar o leite

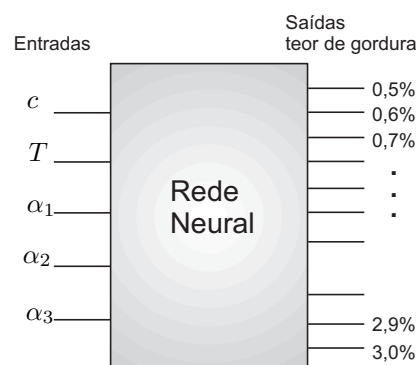


Figura 5.18. Esquema da rede neural projetada para determinação do teor de gordura, com cinco entradas (velocidade de propagação, temperatura, coeficiente de atenuação em três frequências) e 26 saídas, com resolução igual a 0,1%.

UHT e comparar os resultados obtidos por ultra-som com os métodos convencionais. Para operação com temperatura constante foram obtidos bons resultados para a medição de densidade em amostras com adição de água. No entanto, para situações em que a célula de medição foi submetida a bruscas variações de temperatura e a diversos ciclos de temperatura, notou-se que a mesma não forneceu medições precisas para a densidade. Esse fato pode ser explicado em consequência da célula de medição não ser adequada para situações de bruscas variações de temperatura e sim para trabalhar em temperatura aproximadamente constante. Sendo assim, a densidade foi um parâmetro que não pôde ser utilizada para o treinamento das redes neurais, pois foi penalizado pelas variações bruscas de temperatura a qual a célula de medição foi submetida.

Uma outra rede neural foi projetada para detectar adulteração de amostras de leite por adição de água na faixa de 1% até 60%, com uma resolução de 1% entre 1 e 10% de água adicionada, e resolução de 5% entre 10 e 60% de água adicionada. Essa resolução foi utilizada inicialmente considerando a resolução de alguns equipamentos comerciais. Para esse problema, foram utilizadas arquiteturas com 5 nodos na camada de entrada assim como a configuração utilizada para obtenção da quantidade de gordura, 21 nodos na camada de saída e uma camada oculta variando o número de nodos em 26, 40 e 52. Na figura 5.19 apresenta-se a configuração de saída para a rede projetada para detecção de água adicionada ao leite.

A velocidade de propagação da onda (c), o coeficiente de atenuação da onda (α), o teor de gordura do leite UHT (ϕ_v) e a quantidade de água adicionada ao leite em função de uma variação de temperatura entre 25°C e 35°C foram utilizados como dados para o treinamento das redes neurais projetadas. Essa faixa de temperatura foi utilizada, pois observou-se uma maior facilidade na classificação das amostras de leite. Em laticínios os experimentos convencionais são realizados em temperatura

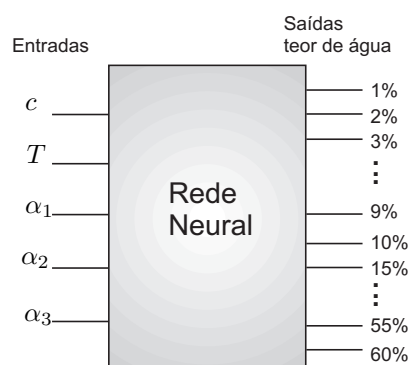


Figura 5.19. Esquema da rede neural projetada para determinação do teor de água, com cinco entradas (velocidade de propagação, temperatura, coeficiente de atenuação em três frequências) e 21 saídas, com resolução igual a 1% (entre 1 e 10%) e 5% (entre 10 e 60%).

ambiente na mesma faixa de temperatura do experimentos realizados no laboratório de ultra-som entre 25°C e 35°C.

A partir da criação de uma rede neural, treinamento e fase de teste, obtém-se um algoritmo capaz de caracterizar os tipos de leite propostos (UHT) e a quantidade de água adicionada ao leite (UHT). Foi utilizado o algoritmo de treinamento por retropropagação (*backpropagation*) com aplicação de um gradiente para o treinamento da rede neural.

5.6.1 Planejamento dos experimentos

No treinamento de uma rede MLP, um dos problemas a serem tratados é a definição dos parâmetros da rede. Essa seleção no entanto é um processo às vezes pouco compreendido, pois há situações na qual podem-se determinar soluções diferentes para o mesmo problema com uma mesma eficiência, e com parâmetros diferentes. Nesse trabalho, variou-se os seguintes parâmetros:

1. Quantidade máxima de iterações (ou ciclos);
2. Quantidade de neurônios na camada intermediária da rede;
3. Taxa de aprendizado;
4. Constante de momento.

Utilizando esses quatro parâmetros projetam-se redes neurais com configurações diferentes, ou seja, para cada combinação dos parâmetros projeta-se uma rede neural de características próprias.

Como critério de parada para realizar os treinamentos da rede, utiliza-se:

- Número máximo de iterações (épocas);
- gradiente mínimo em 10^{-10} ;
- um menor SSE de validação.

Para cada uma das topologias projetadas, foram feitas 10 execuções, ou seja, 10 redes com inicializações diferentes de pesos foram treinadas e a partir disso foram computadas as somas dos quadrado dos erros (SSE) para os conjuntos de treinamento, de validação e de teste. A melhor rede foi escolhida pelo menor valor médio das SSE observadas no conjunto de validação e pelo menor erro de classificação.

5.6.2 Procedimento e descrição dos parâmetros

O problema abordado enquadra-se na classificação de padrões, ou seja, ao se apresentar um determinado padrão, a rede deve ter como resposta a classe a qual ele pertence. Quando o sensor de ultra-som adquirir determinados padrões ou respostas para uma determinada amostra de leite, a rede deve saber a qual tipo de leite esse padrão pertence. A metodologia utilizada faz parte de um relatório técnico para problemas de treinamento de redes neurais [32].

Ao adquirir os dados de temperatura, velocidade e atenuação, é realizado um pré-processamento, onde os dados são normalizados para um intervalo entre 0 e 1. Utiliza-se assim a função sigmoïdal. Para os atributos categóricos e saídas utiliza-se um artifício de codificação ortogonal. Atribui-se uma seqüência de bits ao atributo, sendo que apenas um dos bits vale 1, indicando a categoria à qual a saída pertence. Por exemplo, se um atributo tiver como valores possíveis leite *A* (3,0% t.g.), leite *B* (2,9% t.g.) e leite *C* (0,7% t.g.), leite *D* (0,6 % t.g.) e leite *E* (0,5 % t.g.), então a categoria leite *A* fica codificada como [0...0 0 1]; a categoria *B* fica codificada como [0...0 1 0]; e assim por diante [17], [18], [32]. Esse mesmo pré-processamento foi utilizado para as redes neurais que detectam a quantidade de água adicionada ao leite. Para chegar nas amostras pré-processadas usa-se a Equação 5.2.

$$\text{amostraPre} = \frac{\text{amostra} - \text{amostraMin}}{\text{amostraMax} - \text{amostraMin}}, \quad (5.2)$$

sendo amostraPre a amostra final pré-processada com valores entre 0 e 1, amostraMax o maior valor dentro do conjunto de amostras coletadas e amostraMin o menor valor coletado para o conjunto de cada variável de entrada.

Em seguida, o particionamento de dados adotado é o sugerido em [32], onde 50% dos padrões de cada classe são escolhidos aleatoriamente para treinamento, 25% para validação e 25% para teste. É importante que as proporções entre as classes no conjunto completo sejam mantidas nos conjuntos de treinamento, validação e teste.

Para os experimentos para detecção de quantidade de gordura do leite, foram utilizadas 11700 amostras para treinamento, 5850 amostras para validação e 5850 para testes, já para o problema de detecção de água adicionada utilizaram-se 9450 amostras para treinamento, 4725 para validação e 4725 para teste.

No passo seguinte, tem-se a determinação da topologia da rede, ou seja, a quantidade de atributos de entrada, número de nodos de saída, que é a quantidade de classes a serem classificadas, quantidade de nodos na camada oculta e a quantidade de camadas escondidas.

Utiliza-se a regra de classificação *winner-takes-all*: o nodo de saída que gerar a maior saída define a classe do padrão. A definição da função logística é outro parâmetro importante para o projeto da rede neural. Essa função é utilizada para incluir uma não-linearidade para que a relação de entrada e saída da rede não esteja

reduzida àquela de um perceptron de camada única [30]. A função utilizada é a função de ativação sigmóide logística. Um outro passo importante é a definição do algoritmo de treinamento, pois esse será responsável pelo processo de aprendizagem da rede neural.

A seguir, para finalizar o treinamento utiliza-se uma técnica de poda da rede, conhecida por parada por erro mínimo de validação. É recomendável que o treinamento seja interrompido quando o erro no conjunto de validação atingir um mínimo. A partir deste ponto, supõe-se que a rede só aprenderia detalhes irrelevantes do conjunto de treinamento. O erro para dados de treinamento seria cada vez menor, mas o erro para dados novos (validação) seria cada vez mais alto. Outra técnica utilizada para poda de redes é máximo de iterações permitidas.

Além de variar a quantidade de nodos na camada oculta, variou-se a taxa de aprendizado das RNAs em 0,01 0,001 e 0,0001 e a quantidade máxima de iterações em 50000, 100000, 150000 e 200000 para o problema de quantidade de gordura e 20000, 40000, 50000 e 100000 para detecção de água no leite. Utilizou-se uma taxa de momento de 0,0002, 0,0008 e 0,0001.

O dimensionamento das redes, como pode ser observado, foi realizado através do método de refinamento manual, estudando os parâmetros e escolhendo a melhor combinação testada. Não foi estudada uma dependência estatística entre esses dados. Para isso poderia-se ter aplicado um outro método, por exemplo o de análise de variância [30].

5.6.3 Execução dos experimentos

Utilizou-se o *software* Matlab, com um pacote de ferramentas específico, para desenvolver as redes neurais que classificam o tipo de leite e a quantidade de água adicionada através de valores de atenuação e velocidade de propagação da onda pela amostra em função da temperatura, adquiridos da célula de medição. Foram testadas Redes MLP completamente conectadas. As redes neurais foram constituídas por três camadas: a primeira foi formada pelo vetor de características previamente selecionado; seguido de uma camada intermediária; por fim a camada de saída que foi constituída por vinte e seis neurônios para a obtenção de quantidade de gordura e vinte e um neurônios para a detecção de água adicionada.

5.6.4 Resultados

Foram treinadas as redes MLP e computou-se a SSE no conjunto de treinamento e no conjunto de validação com execuções para cada combinação de parâmetros. Em cada execução tem-se uma inicialização de pesos diferente, e sabe-se que obtêm-se resultados diferentes para cada execução. Para cada topologia usa-se a mesma inicialização de pesos quando varia-se a taxa de treinamento e o número máximo

Tabela 5.7. Redes treinadas para a topologia com três camadas de entrada para detecção do teor de gordura.

Nodos	Iterações	SSE treinamento	SSE validação	SSE teste	Erro
26	50000	5060,0	2595,0	2594,3	28,3%
26	100000	3630,0	1887,0	1893,2	19,1%
26	150000	3490,0	1806,0	1826,5	16,9%
26	200000	3224,3	1600,2	1650,5	14,5%
40	50000	4050,0	2075,0	2089,3	20,8%
40	100000	3280,0	1722,0	1726,0	17,5%
40	150000	3288,6	1704,0	1728,8	15,3%
40	200000	3040,3	1586,3	1586,4	10,2%
52	50000	4119,0	2130,4	2143,7	20,4%
52	100000	3106,0	1633,0	1643,8	17,3%
52	150000	3048,9	1601,0	1609,6	14,8%
52	200000	2389,1	1236,0	1249,9	8,0%

de iterações e, então, o algoritmo projetado salva a melhor rede, ou seja, a que apresentou a menor soma de quadrados dos erros (SSE) no conjunto de validação e a menor percentagem de erro.

Desta forma foram selecionadas as redes neurais projetadas com topologias diferentes e maior capacidade de generalização. A Tabela 5.7 apresenta os resultados para a topologia composta por três entradas. Na primeira e segunda colunas encontram-se o número de nodos na camada oculta (Nodos) e o número máximo de iterações (Iterações), respectivamente. As colunas seguintes são relativas aos SSEs dos conjuntos de treinamento (SSE treinamento), validação (SSE validação) e teste (SSE teste), respectivamente. A última coluna é relativa ao erro de classificação das redes (Erro). A taxa de aprendizado η utilizada foi de 0,0001. Inicialmente foram realizados testes com taxas de aprendizado de 0,01 e 0,001, mas o treinamento apresentou uma subida no erro de validação maior que a determinada, sendo assim os treinamentos foram finalizados de forma precoce, levando as redes a uma não generalização do problema. Para estas configurações utilizou-se uma constante de momento de 0,0008.

Além das SSEs, computou-se o erro de classificação das redes. Observa-se que, para as configurações apresentadas na Tabela 5.7 o menor erro de classificação no conjunto de teste foi de 8%, o que representa 92% de classificação correta no teor

Tabela 5.8. Redes treinadas para a topologia com cinco camadas de entrada para detecção do teor de gordura.

Nodos	Iterações	SSE treinamento	SSE validação	SSE teste	Erro
26	50000	4793,9	2441,5	2463,8	26,4%
26	100000	2226,5	1150,0	1163,5	9,6%
26	150000	1893,4	987,5	995,8	8,3%
26	200000	1835,2	937,9	965,0	8,2%
40	50000	4380,8	2242,0	2274,6	22,1%
40	100000	2025,1	1040,6	1662,5	8,8%
40	150000	1520,3	792,5	812,2	10,3%
40	200000	1423,9	755,1	762,1	6,3%
52	50000	4025,0	2057,4	2092,7	19,1%
52	100000	2684,1	1406,2	727,0	12,4%
52	150000	2064,5	1095,9	1114,6	6,2%
52	200000	1371,1	724,5	738,0	4,5%

de gordura das amostras analisadas.

Verifica-se que a rede que apresentou, em média, o menor SSE do conjunto de validação foi a que contém 52 nodos na camada oculta e 200000 iterações. Esta rede apresentou um erro de classificação no conjunto de teste de 8%.

Para a topologia composta com 5 entradas foram incorporadas ao algoritmo de treinamento dados de atenuação na faixa de 4 MHz e 6,5 MHz. Para essa configuração também foram usadas as taxas de aprendizado de 0,01 e 0,001, mas novamente o treinamento foi finalizado de forma precoce. Notou-se para estes casos que o algoritmo não consegue ter uma faixa de classificação precisa. Utilizou-se então uma taxa de aprendizado de η de 0,0001 e uma constante de momento α de 0,0001. A Tabela 5.8 apresenta os resultados para as melhores configurações utilizadas.

Nota-se que os experimentos que utilizam cinco entradas tiveram melhor classificação das amostras utilizadas no grupo de teste que a topologia formada por três camadas. Isso mostra que aumentando a quantidade de dados, que forneçam representatividade da população das amostras usadas, pode ser obter redes com melhor índice de classificação, ou seja, maior precisão. Para a melhor configuração treinada, pode-se observar uma classificação correta de 95,5% e um erro de 4,5% para a rede com 52 nodos na camada oculta, 200000 iterações e constante de momento de 0,0001. O tempo de treinamento para essa topologia foi muito longo e isso não se explica

somente pela quantidade maior de iterações mas também pela maior complexidade dos cálculos fornecida pelo aumento na quantidade de nodos escondidos. Na Figura 5.20 apresenta-se uma comparação entre valores de classificação correta para redes com três e cinco nodos na camada de entrada, 26, 40 e 52 nodos na camada oculta e 200000 iterações. Ela mostra que tem-se uma grande proximidade na capacidade de classificação correta. Como não foi encontrado a incidência de *overfitting* e o objetivo foi de alocar para a rede um número de nodos ocultos suficientes para a solução do problema a rede utilizada para classificar a quantidade de gordura no leite foi a com 5 nodos na camada de entrada e 52 nodos na camada oculta.

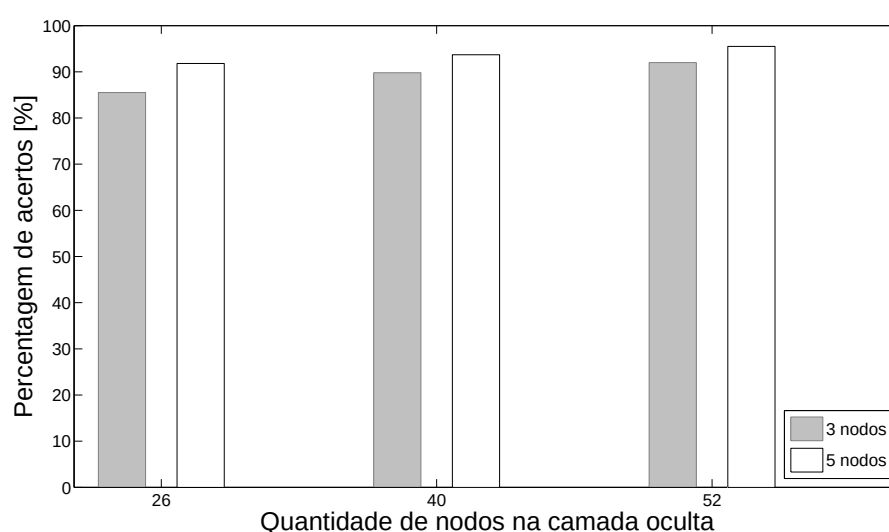


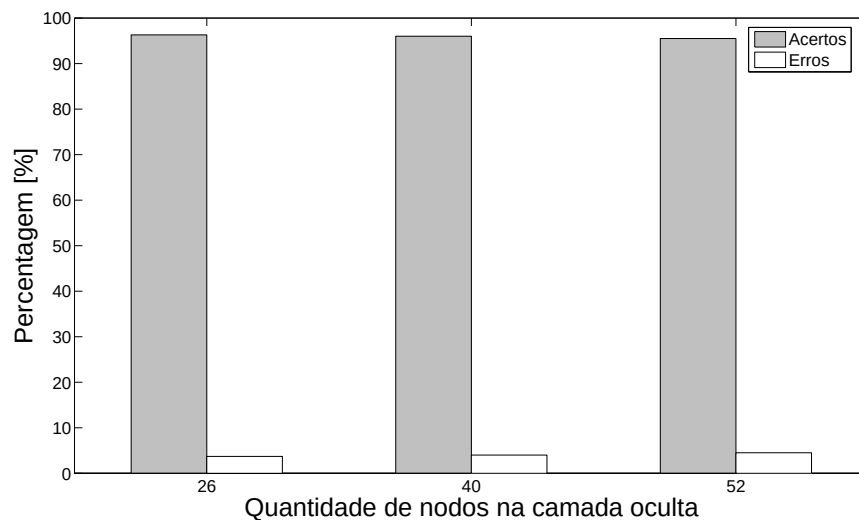
Figura 5.20. Capacidade de classificação usando três e cinco nodos na camada de entrada para detecção do teor de gordura.

A Tabela 5.9 apresenta os resultados obtidos para o problema de detecção de água adicionada. Observa-se que foram obtidos resultados com grande capacidade de classificação. As configurações com 40 e 52 nodos na camada oculta obtiveram o maior potencial de classificação. Para a topologia com 40 nodos obteve-se um erro de 4,3%, sendo assim uma classificação correta de 95,7%. Para a topologia com 52 nodos obteve-se uma classificação correta de 96,3% e um erro de 3,7% sendo que para esta rede o tempo de treinamento foi maior que para a configuração com 40 nodos ocultos, isso provavelmente pelo aumento na complexidade dos cálculos envolvidos nas operações.

Na Figura 5.21 apresenta-se uma relação entre a capacidade de classificação entre as configurações com 26, 40 e 52 nodos na camada oculta e 5 nodos na camada de entrada. Essa Figura (5.21) mostra que para o problema de água adicionada conseguiu-se resultados muito próximos de classificação correta das amostras entre as diversas configurações testadas, o que torna qualquer uma das configurações apresentadas capaz de classificar com precisão a quantidade de água adicionada ao leite.

Tabela 5.9. Redes treinadas para a topologia com cinco camadas de entrada para detecção de água adicionada.

Nodos	Iterações	SSE treinamento	SSE validação	SSE teste	Erro
26	10000	4552	2309	2266,5	22,1%
26	40000	2898	1456	1470,6	16,4%
26	50000	2428	1236	1204,2	12,9%
26	100000	931,1	514,9	485,0	4,5%
40	10000	3900	1981	1940,9	17,6%
40	40000	2025	1031	1023	9,4%
40	50000	1553	795	763,7	7,2%
40	100000	870,4	470,7	469,9	4,0%
52	10000	3834	1950	1908	18,2%
52	40000	1778	904	879,2	9,2%
52	50000	1544	753,9	774,2	7,6%
52	100000	752,5	425,7	403,1	3,7%

**Figura 5.21.** Capacidade de classificação usando cinco nodos na camada de entrada para detecção da quantidade de água adicionada ao leite.

A soma de quadrados dos erros SSE é incluído nas Tabelas 5.7, 5.8 e 5.9 somente para efeito de registro e como critério de parada para o treinamento das redes neurais, uma vez que um erro médio quadrado pequeno ou grande não implica em uma generalização correta. A capacidade de classificação correta é o principal parâmetro

para escolha das melhores redes neurais projetadas.

Pode-se observar na Figura 5.22 o decaimento do erro para a rede MLP para detecção do teor de gordura com 52 nodos na camada oculta, taxa de aprendizado 0,0001 e máximo de 200000 iterações, que os erros para os conjuntos de validação e treinamento tendem a decrescer. Nesse gráfico não há evidências de que ocorre um *overfitting* (quando no gráfico do erro os dados de validação e treinamento se cruzam), devido ao fato de que o treinamento da rede alcançou o número máximo de iterações sem que o erro de validação tenha subido por 3 iterações consecutivas, isso garante que a rede pode estar generalizando de forma satisfatória produzindo resultados precisos de classificação das amostras.

Utiliza-se a capacidade de classificação como parâmetro para a escolha da melhor topologia e essa é usada por um outro algoritmo que testa qual a quantidade de gordura presente ao leite. Da mesma forma escolheu-se uma rede para detectar a quantidade de água adicionada ao leite. A partir desses algoritmos pode-se carregar um novo conjunto de teste para caracterizar o leite.

Foram realizados testes com dados de experimentos com amostras de leite UHT com teor de gordura e quantidade de água adicionada previamente conhecidos, mas desconhecidos pela rede neural. Esses testes são denominados de testes cegos. As aquisições de dados (temperatura, velocidade e atenuação) foram feitas, os dados armazenados e utilizados para alimentar a entrada da rede neural e assim avaliados pela mesma. Os resultados comprovaram a eficácia do algoritmo que obteve para as duas configurações, referentes a quantidade de gordura e água adicionada ao leite, no geral um potencial na faixa de 98 a 100% de acerto nos testes de quantidade de concentração de gordura e quantidade de água adicionada ao leite UHT.

Utilizou-se uma configuração com a rede neural projetada para detecção do teor de gordura e a rede para detecção de água adicionada de forma simultânea, ou seja, realizou-se o teste de gordura seguido do teste de quantidade de água adicionada. Esse teste é de grande importância, pois constatou-se que uma amostra de leite com uma quantidade qualquer de água adicionada pode ter a mesma velocidade de propagação que um leite de um determinado teor de gordura, mas sem água adicionada. Esse teste é possível pois utilizando redes neurais, como visto anteriormente, pode-se trabalhar com diversas variáveis para obtenção de uma determinada resposta. O algoritmo não depende de uma única variável como velocidade de propagação, mas pode-se usar conjuntamente atenuação, velocidade e temperatura. Essas variáveis desde de que adquiridas de forma confiável são utilizadas no treinamento da rede neural, que posteriormente será utilizada para determinar as características do leite (teor de gordura e quantidade de água adicionada). O projeto de uma rede neural para detecção simultânea do teor de gordura e água adicionada pode exigir um grande esforço computacional, pois seria adicionado uma grande quantidade de dados, nodos na camada oculta e de saída, etc.

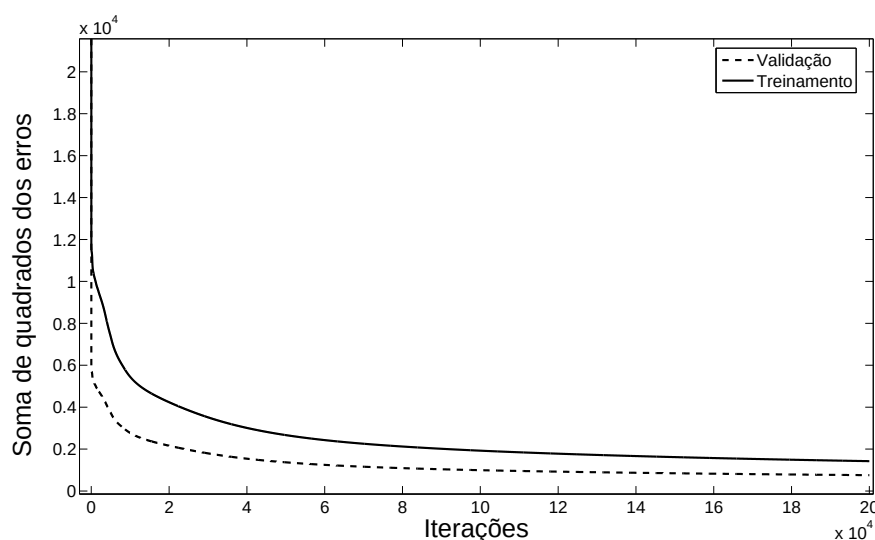


Figura 5.22. Decaimento do erro de treinamento (linha sólida) e validação (linha tracejada), para detecção do teor de gordura.

5.6.5 Dificuldades encontradas e limitações do modelo

As redes MLP que utilizam o algoritmo *backpropagation* no aprendizado, assim como muitos outros tipos de RNA, diferentemente dos modelos probabilísticos, podem ser vistas como “caixas pretas”, uma vez que os modelos não apresentam justificativas para suas respostas. Neste contexto, diversos pesquisadores vêm trabalhando em busca de extração de conhecimento de RNAs, e na criação de procedimentos explicativos, onde se tenta justificar o comportamento da rede em determinadas situações.

Uma dificuldade do modelo MLP treinado por *backpropagation* refere-se ao custo computacional (tempo de treinamento), que tende a ser muito alto. Algumas vezes são necessários milhares de iterações para se alcançar níveis de erros aceitáveis, pois as funções são calculadas para cada unidade e suas conexões separadamente, o que pode ser problemático em redes muito grandes, ou com grande quantidade de dados.

Em particular, é muito complexo decidir qual a arquitetura ideal de uma rede MLP de maneira que ela seja tão grande quanto o necessário para conseguir obter as representações necessárias, ao mesmo tempo pequena o suficiente para se reduzir o tempo de treinamento. Outra dificuldade encontrada é o fato de não existirem regras claras para se definir quantas unidades devem existir nas camadas intermediárias, quantas camadas, ou como devem ser as conexões entre essas unidades.

5.7 Comentários

Foram realizados experimentos com leite UHT integral e desnatado, variando-se a concentração de gordura, teor de água adicionada e temperatura. Para obter

as amostras com variação de 0,1% de concentração de gordura foram realizados diluições entre o leite UHT, com 3% de teor de gordura, e o leite desnatado, com 0,5%. As diluições de água destilada nas amostras de leite UHT foram feitas numa proporção entre 1% a 60% como o objetivo de detectar a adulteração das amostras.

A preparação das amostras foi uma etapa de extrema importância para os resultados da pesquisa. Essa preparação consta de um processo bastante cuidadoso e detalhado, desde a pesagem das amostras, para se obter concentrações corretas de gordura e água adicionada para cada amostra a ser analisada, até a verificação de variação de temperatura e o manuseio com os equipamentos. As amostras de leite UHT passam por um processo de calibração, para obter os valores do teor de gordura e água adicionada. Para isso, foram realizados experimentos pelos métodos de Gerber e crioscópio, o que demandou extremos cuidados na diluição e homogeneização das amostras para evitar resultados inconsistentes.

Os resultados para a variação da densidade em função do teor de gordura não apresentaram poder de discriminação, pois os valores sofrem pequenas variações estando dentro do erro de medida. Observa-se que a densidade pode ser utilizada com maior poder de discriminação para amostras de leite com adição de água. Assim, a densidade torna-se uma variável importante para inibir possíveis fraudes por adição de água. Essa variável não foi utilizada para o treinamento de redes neurais utilizadas para detecção de quantidade de água adicionada, pois a célula de medição sofreu com o problema de variações bruscas de temperatura que causaram medições imprecisas.

Os experimentos por ultra-som com variação de temperatura demandam muito tempo em função da transferência de calor entre o sistema (banho ultratermostático) e as partes que formam a célula. Soma-se a este fato a necessidade de se manter a temperatura estabilizada em determinadas faixas. A variação de temperatura foi de 15°C a 35°C. Para cada amostra de leite o tempo médio para esses experimentos foi de 8 a 10 horas e foram realizados vários experimentos para validar os resultados. Na prática pode-se operar com temperatura constante. Utilizando técnicas de redes neurais pode-se trabalhar com uma variação de temperatura, pois essa variável pode ser inserida ao treinamento levando a rede a generalizar e reconhecer novos dados inseridos experimentalmente.

Para o treinamento das redes neurais projetadas foi realizado primeiro o pré-processamento das amostras. Como o número de amostras foi muito grande, esse processo também demandou um enorme tempo. Na fase de treinamento, foram testadas várias configurações de redes para se chegar em uma configuração adequada para a resolução do problema de classificação de teor de gordura no leite e da quantidade de água adicionada ao leite. Assim, depois do treinamento de diversas redes neurais foram escolhidas algumas configurações que puderam ser utilizadas para caracterizar as amostras corretamente. O treinamento de algumas redes foi realizado

durante vários dias, devido à enorme quantidade de dados.

Observa-se que a utilização de diversos parâmetros pode melhorar a capacidade de decisão das redes neurais. Por exemplo, num caso no qual a concentração pode variar com o teor de gordura do leite e com a quantidade de água adicionada, a mesma velocidade de propagação pode representar duas concentrações diferentes, e o uso de diversos parâmetros pode reduzir a incerteza no poder de classificação da rede neural.

Capítulo 6

Conclusões

O trabalho apresentou um aspecto multidisciplinar em que foram abordados temas como leite e técnicas de análises convencionais (químico e físicos), métodos de caracterização de líquidos por ultra-som e métodos de classificação de padrões utilizando redes neurais. Este foi o primeiro trabalho do grupo de pesquisa utilizando redes neurais.

Utilizou-se uma célula de medição para obter dados de densidade, velocidade de propagação e coeficiente de atenuação acústicos, que foram relacionados com a concentração de gordura e água adicionada, medidos também com métodos convencionalmente utilizados em laticínios, para efeito de calibração de amostras e posterior utilização pelas redes neurais artificiais.

A velocidade é um parâmetro bastante sensível a qualquer um dos parâmetros (teor de gordura e água adicionada), mas que também depende fortemente da temperatura. Sendo assim, experimentos com temperatura controlada são importantes para se obter resultados confiáveis. A atenuação foi um parâmetro que não sofreu alterações significativas com a variação de temperatura, o que a torna muito importante para o método, apesar de não apresentar a mesma resolução que a velocidade. A densidade foi um parâmetro com o qual não se obtiveram bons resultados para discriminar o teor de gordura, mas apresentou boa capacidade para discriminar adição de água. No entanto, problemas oriundos da operação da célula de medição, em ciclos repetidos de temperatura, provocaram o mal-funcionamento desta para a medição da densidade, o que levou a descartar o uso deste parâmetro na caracterização das amostras. Salienta-se que o uso da densidade pode ser muito importante na discriminação de água, principalmente.

Foram utilizadas redes neurais, que através de simulações e treinamento foram utilizadas para a resolução dos problemas de classificação do teor de gordura e de quantidade de água adicionada ao leite. A partir dos vários treinamentos realizados com diferentes topologias e parâmetros, foram escolhidas algumas redes para classificar as amostras de leite, com concentrações de gordura entre 0,5% e 3,0% e quantidade de água adicionada entre 1% e 60%. Essas redes foram utilizadas pelo

algoritmo, que diagnosticou com sucesso o tipo de leite, pela concentração de gordura e pela quantidade de água adicionada, numa faixa de temperatura entre 25 e 35°C.

As técnicas utilizadas permitiram obter uma resolução de 0,1% na determinação do teor de gordura. Para a determinação do teor de água adicionada (a.a.), a resolução foi de 1% para a faixa entre 1% e 10% a.a. e de 5% para amostras entre 10% e 60% a.a. Essas faixas de resolução mostram que os métodos por ultra-som e redes neurais podem ser equivalentes aos métodos de Gerber, ponto de congelamento e equipamentos comerciais disponíveis no mercado.

Observou-se que as redes neurais possibilitaram grande poder de classificação e isso foi comprovado nos experimentos de testes com amostras desconhecidas pelas redes projetadas. Para as redes projetadas para o problema de adição de água obteve-se um poder de classificação de 96,3% e para a obtenção do teor de gordura 95,5% de classificação correta.

As melhores taxas de treinamento e momento utilizadas foram 0,0001 e 0,0008, respectivamente. A utilização de apenas uma camada oculta foi suficiente para que a rede neural generalizasse bem, e isso foi analisado através dos experimentos realizados.

Uma das principais contribuições deste trabalho foi mostrar uma sequência de procedimentos e experimentos para encontrar uma rede neural que generalizasse bem e pudesse ser utilizada em problemas de reconhecimento de padrões. Outra contribuição foi de validar os métodos de ultra-som e redes neurais para caracterização de quantidade de gordura e água adicionada ao leite UHT.

Os métodos de ultra-som e redes neurais mostram-se eficazes para o problema proposto e possuem várias vantagens em relação a métodos convencionais, como rapidez (uma vez treinada a rede), são métodos não-destrutivos e podem ser não-agressivos.

Nota-se que foram obtidos resultados mais precisos para medição de velocidade de propagação em comparação aos resultados de atenuação. Isso deve-se ao fato que a célula de medição inicialmente foi projetada para medição de densidade e velocidade de propagação da onda. As técnicas de processamento digital de sinais possibilitam a medição da velocidade de propagação de forma confiável e precisa. A medição de atenuação, no momento, é mais complicada devido a geometria da célula de medição e da própria constituição do leite. A utilização de uma nova célula de medição otimizada e com maior estabilidade para a medição de atenuação e densidade pode melhorar os resultados de classificação das amostras de leite.

Para trabalhos futuros sugere-se o desenvolvimento de uma nova célula de medição com uma geometria mais robusta e compacta, que possibilite maior resolução e precisão para medidas de velocidade e atenuação acústicas. Outro ponto de grande importância é a utilização desta nova célula para estudar as características acús-

licas do leite *in natura*, que apresenta dificuldades em relação à homogeneidade e presença de glóbulos de gordura de diferentes dimensões. Sugere-se ainda o estudo de novas configurações de redes neurais para a caracterização de leite e ainda a implementação das mesmas em dispositivos como FPGA e DSPs com o objetivo de desenvolver equipamentos para caracterização do leite *in natura*. Outro ponto importante é a necessidade do estudo envolvendo o tipo de leite *in natura*, gado, alimentação e região de criação do gado leiteiro, pois esses fatores influenciam na caracterização do leite. Em decorrência da pesquisa e dos resultados obtidos, foram apresentados trabalhos no 5th Brazilian Conference on Dynamics, Control and Their Applications, em 2006, Guatatingueta, SP [37]. Congresso Brasileiro de Ciência e Tecnologia de alimentos, em 2006, Curitiba, PR [38]. Congresso Brasileiro de Autômática, em 2006, Salvador, BA [39]. International Congress on Ultrasonics, em 2007, Viena, Austria [40].

Referências Bibliográficas

- [1] YAMAGUCHI, L. C. T.; MARTINS, P. C.; CARNEIRO, A. V. Análise da evolução do agronegócio do leite no Brasil: 1970-1999. In: CONGRESSO BRASILEIRO DE ECONOMIA E SOCIOLOGIA RURAL, 40., 2002, Passo Fundo. **Anais...** Brasília: Sociedade Brasileira de Economia, Administração e Sociologia Rural, 2002. p. 1–16.
- [2] MARTINS, P. C.; YAMAGUCHI, L. C. T.; CARNEIRO, A. V. Por que calcular o custo da produção do leite. **Informe econômico do leite**, Juiz de Fora, v. 3, n. 3, 2003.
- [3] CARVALHO, M. P. Perspectivas brasileiras no comércio mundial de lácteos. In: ANUÁRIO de pecuária brasileira. São Paulo: FNP, 2004, p. 196–200.
- [4] BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Instrução Normativa n. 51 de 18 de setembro de 2002. **Diário Oficial da União - Imprensa Nacional**, seção 1, p. 13, 2002.
- [5] BEHMER, M. L. A. **Tecnologia do leite**: produção, industrialização e análise. 13. ed. São Paulo: Nobel, 1984. 320 p.
- [6] POVEY, M. J. W. **Ultrasonic techniques for fluids characterization**. San Diego: Addison-Wesley, 1997. 214 p.
- [7] MCCLEMENTS, D. J. Advances in the application of ultrasound in food analysis and processing. **Trends in Food Science Technology**, Amsterdam, v. 6, n. 9, p. 293–299, 1995.
- [8] MASON, T. J.; PANIWNKY, L.; LORIMER, J. The uses of ultrasound in food technology. **Ultrasonics Sonochemistry**, Amsterdam, v. 3, n. 3, p. 253–260, 1996.
- [9] JAVANAUD, C. Applications of ultrasound to food systems. **Ultrasonics**, Amsterdam, v. 26, n. 3, p. 117–123, 1988.
- [10] BACANELI, F. **Determinação do teor de gordura no leite através de parâmetros acústicos**. 1998. 106 f. Dissertação (Mestrado em Engenharia) - Escola Politécnica, Universidade de São Paulo, São Paulo, 1998.

-
- [11] DORABIATO, L. F.; MAIA, J. M.; GAMBA, H. R. Caracterização de leite bovino por ultra-som. In: CONGRESSO BRASILEIRO DE ENGENHARIA BIOMÉDICA, 20., 2006, São Pedro. **Anais...** Campinas: Sociedade Brasileira de Engenharia Biomédica, 2006. p. 788–791.
- [12] GUNASEKARAN, S.; CHYUNG., A. Evaluating milk coagulation with ultrasonics. **Food technology**, Chicago, v. 48, n. 12, p. 74–78, 1994.
- [13] HIGUTI, R. T.. **Caracterização de líquidos por ultra-som**. 2001. 127 f. Tese (Doutorado em Engenharia) - Escola Politécnica, Universidade de São Paulo, São Paulo, 2001.
- [14] ADAMOWSKI, J. C. **Medição de densidade de líquidos por ultra-som**. 1993. 98 f. Tese (Doutorado em Engenharia) - Escola Politécnica, Universidade de São Paulo, São Paulo, 1993.
- [15] MILES, C. A.; SHORE, D.; LANGLEY, K. R. Attenuation of ultrasound in milks and creams. **Ultrasonics**, Amsterdam, v. 28, n. 6, p. 394–400, 1990.
- [16] BRAGA, A. P.; CARVALHO, A. C. L. F.; LUDEMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. Rio de Janeiro: LTC, 2000. 262 p.
- [17] YAN, H.; JIANG, Y.; ZHENG, J.; PENG, C.; LI, Q. A multilayer perceptron-based medical decision support system for heart disease diagnosis. **Expert Systems with Applications**, Amsterdam, v. 30, n. 2, p. 272–281, 2006.
- [18] ERGUN, U.; SERHATLIOGLU, S.; HARDALAÇ, F.; GULER, I. Classification of carotid artery stenosis of patients with diabetes by neural network and logistic regression. **Computers in Biology and Medicine**, Amsterdam, v. 46, n. 1, p. 389–405, 2004.
- [19] TIAN, H.; SHANG, Z. Artificial neural network as a classification method of mice by their calls. **Ultrasonics**, Amsterdam, v. 44, Suppl. 1, p. e275–e278, 2006.
- [20] OCHOA, A.; HERNÁNDEZ, A.; UREÑA, J.; MAZO, M.; JIMENEZ, J. A.; GARCÍA, J.; JIMÉNEZ, A. Ultrasonic pattern recognition based on ANN and PCA techniques. In: IEEE INTERNATIONAL SYMPOSIUM ON INDUSTRIAL ELECTRONICS, 22., 2005, Dubrovnik, Croácia. **Proceedings...** Dubrovnik: IEEE Press, 2005. p. 1215–1220.
- [21] JIAN, X.; GUO, N.; DU, H.; LI, M. X.; ZHIANG, H. L. Automatic defect characterization using artificial neural networks and deconvolution techniques. **Intelligent Material Systems and Structures**, London, v. 17, p. 713–720, Aug/Sep 2006.

-
- [22] ALAIS, C. **Ciencia de la Leche**: principios de técnica lechera, 2. ed. México: CIA Editorial Continental, 1970. 348 p.
- [23] KINSLEY, L. E.; FREY, A. R.; COPPENS, A. B.; SANDERS, J. V. **Fundamentals of Acoustics**. 3. ed. Singapore: John Wiley Interscience, 1982. 496 p.
- [24] EPSTEIN, P. S.; CARHART, R. R. The absorption of sound in suspensions and emulsions. **J. Acoust. Soc. Am.**, Melville, v. 25, n. 3, p. 553–565, 1953.
- [25] DUKHIN, A.; GOETZ, P. J. **Ultrasound for characterizing liquid-based food products**. Dispersion Technology: Bedford Hills, (2002?). Disponível em: < [www.dispersion.com / pages / newsletter / articles / Newsletter14.htm](http://www.dispersion.com/pages/newsletter/articles/Newsletter14.htm) >. Acesso em 20 abr. 2006.
- [26] ALLEGRA, A. S. Effect of particle velocity on propagation of sound in suspensions and emulsions. **J. Acoust. Soc. Am.**, Melville, v. 51, n. 1, p. 182–191, 1972.
- [27] FUKUNAGA, K. J. **Introduction to statistical pattern recognition**. 2. ed. New York: Academic Press, 1990. 592 p.
- [28] MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **Bulletin of mathematical biophysics**, New York, v. 5, p. 115–133, 1943.
- [29] WIDROW, B.; LEHR, M. A. 30 years of adaptive neural networks: perceptron, adaline, and backpropagation. **Proceedings of the IEEE**, New Jersey, v. 78, n. 9, p. 1415–1442, 1990.
- [30] HAYKIN, S. **Redes neurais**: princípios e práticas. 2. ed. Porto Alegre: Bookman, 2001. 898 p.
- [31] KNIGHT, K. Connection ideas and algorithms. **Communications of the ACM**, New York, v. 33, n. 11, p. 59–74, 1990.
- [32] PRECHELT, L. **A set of neural network benchmark problems and benchmarking rules**. Karlsruhe: Universitat Karlsruhe, 1994. v.21. Disponível em: < <http://citeseer.ist.psu.edu/prechelt94proben.html> >. Acesso em 18 mai. 2006.
- [33] RUSSEL, S.; NORVIG, P. **Inteligência artificial**. São Paulo: Editora Campus-Elsevier, 2004. 1056 p.
- [34] DEL-GROSSO, V. A.; MADER, C. W. Speed of sound in pure water. **J. Acoust. Soc. Am.**, Melville, v. 52, n. 5B, p. 1442–1446, 1972.

-
- [35] MILLER, R. M. **Flow measurement engineering handbook**. 2. ed. New York: McGraw Hill, 1988. 1168 p.
- [36] HIGUTI, R. T.; GALINDO, B. S.; KITANO, C.; BUIOCHI, F.; ADAMOWSKI, J. C. Thermal characterization of an ultrasonic density-measurement cell. **IEEE Transactions on Instrumentation and Measurement**. New Jersey, v. 56, n. 3, p. 924–930, 2007.
- [37] NAZARIO, S. L. S.; KITANO, C.; DOS SANTOS ISEPON, J.; HIGUTI, R. T. Obtenção do teor de gordura em leite através de técnicas de ultra-som e redes neurais. In: BRAZILIAN CONFERENCE ON DYNAMICS, CONTROL AND THEIR APPLICATIONS, 5., 2006, Guaratinguetá, **Anais...** [S.l.]: [S.n.], 2006. 6p.
- [38] NAZARIO, S. L. S.; DOS SANTOS ISEPON, J.; HIGUTI, R. T. Medição de teor de gordura em leite usando ultra-som. In: CONGRESSO BRASILEIRO DE CIÊNCIA E TECNOLOGIA DE ALIMENTOS, 20., 2006, **Resumos...** Campinas: Sociedade Brasileira de Ciência e Tecnologia de Alimentos, 2006. 1 p.
- [39] NAZARIO, S. L. S.; DE OLIVEIRA, H. C. B.; KITANO, C.; DOS SANTOS ISEPON, J.; HIGUTI, R. T. Classificação do teor de gordura em leite uht utilizando técnicas de ultra-som e redes neurais em função da temperatura. In: CONGRESSO BRASILEIRO DE AUTOMÁTICA, 16., 2006, Salvador. **Anais...** Campinas: Sociedade Brasileira de Automática, 2006. 4 p.
- [40] NAZARIO, S. L. S.; KITANO, C.; HIGUTI, R. T.; ISEPON, J. S.; BUIOCHI, F. Use of ultrasound and neural networks to detect adulteration in milk samples. In: INTERNATIONAL CONGRESS ON ULTRASONICS, 1., 2007, Viena. **Proceedings...** [S.l.]: [S.n.], 2007. 4 p.



,UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
Câmpus de Ilha Solteira
Programa de Pós-Graduação em Engenharia Elétrica
Av. Brasil Centro, 56
15385-000 Ilha Solteira - SP
www.dee.feis.unesp.br