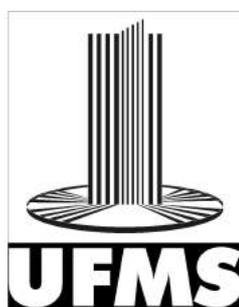


Luiz Fernando Delboni Lomba

Identificação do Comportamento Bovino a partir dos Dados de Movimentação e do Posicionamento do Animal



Campo Grande - MS

Dezembro de 2015

Luiz Fernando Delboni Lomba

Identificação do Comportamento Bovino a partir dos Dados de Movimentação e do Posicionamento do Animal

Trabalho apresentado como requisito para obtenção do título de mestre em Computação Aplicada.

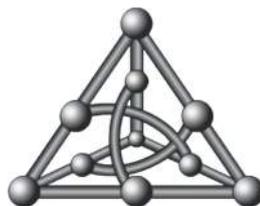
Orientadora: Prof^a. Dr^a. Hana Karina Salles Rubinsztein

Coorientador: Dr. Pedro Paulo Pires

Universidade Federal de Mato Grosso do Sul – UFMS

Faculdade de Computação

Mestrado em Computação Aplicada



Campo Grande - MS

Dezembro de 2015

Luiz Fernando Delboni Lomba

Identificação do Comportamento Bovino a partir dos Dados de Movimentação e do Posicionamento do Animal/ Luiz Fernando Delboni Lomba. – Campo Grande - MS, Dezembro de 2015-

115 p. : il. (algumas color.) ; 30 cm.

Orientadora: Prof^a. Dr^a. Hana Karina Salles Rubinsztein

Coorientador: Dr. Pedro Paulo Pires

Dissertação (Mestrado) – Universidade Federal de Mato Grosso do Sul – UFMS
Faculdade de Computação

Mestrado em Computação Aplicada, Dezembro de 2015.

1. Etologia. 2. Comportamento Animal. 3. Pecuária de Precisão. 4. Reconhecimento de Padrões. 5. Sensores. I. Orientadora Prof^a. Dr^a. Hana Karina Salles Rubinsztein. II. Universidade Federal de Mato Grosso do Sul. III. Faculdade de Computação. IV. Mestrado em Computação Aplicada V. Identificação do Comportamento Bovino a partir dos Dados de Movimentação e do Posicionamento do Animal

Luiz Fernando Delboni Lomba

Identificação do Comportamento Bovino a partir dos Dados de Movimentação e do Posicionamento do Animal

Trabalho apresentado como requisito para obtenção do título de mestre em Computação Aplicada.

Orientadora: Prof^a. Dr^a. Hana Karina Salles Rubinsztein

Coorientador: Dr. Pedro Paulo Pires

Trabalho aprovado. Campo Grande - MS, 10 de dezembro de 2015.

**Prof^a. Dr^a. Hana Karina Salles
Rubinsztein**
Orientadora

Dr. Pedro Paulo Pires
Coorientador

Prof. Dr. Luciano Gonda
Convidado

Prof. Dr. Eraldo R. Fernandes
Convidado

Prof. Dr. Irineu Sotoma
Convidado

Campo Grande - MS
Dezembro de 2015

Agradecimentos

A Deus, pelo dom da vida e pela sabedoria concedida para superar os obstáculos que surgiram no decorrer dos últimos três anos.

Aos meus pais, Luiz e Marlene, que desde sempre me apoiaram e mesmo de longe me incentivaram nesta jornada. Não tenho dúvidas que não chegaria aqui sem a ajuda de vocês.

A professora Hana Karina Salles Rubinsztein, que aceitou o desafio me orientar e que contribuiu em detalhes fundamentais deste trabalho. Obrigado pelo apoio e compreensão em todo este período.

Ao Dr. Pedro Paulo Pires, pela maneira que enxerga as possibilidades e estimula a busca por novos horizontes.

Aos coordenadores do Mestrado Profissional, professores Luciano Gonda e Débora Maria Barroso Paiva, pela atenção e apoio dispensados.

A todos os professores da FACOM, em especial ao professor Irineu, que encabeçou o projeto de fomento que possibilitou a compra dos equipamentos utilizados, e a FUNDECT pelo apoio financeiro dispendido.

A toda a equipe da Embrapa Gado de Corte, em especial aos profissionais do Mangueiro Digital, Quintino Izidio dos Santos Neto, Daniel Oliveira Martins e João Batista Gonçalves Costa Junior, que me auxiliaram e contribuíram na realização e validação dos experimentos em campo.

Aos professores Ederson Marcos Sgarbi e Carlos Eduardo Ribeiro, que me incentivaram desde a graduação, a seguir a carreira acadêmica e me indicaram com as cartas de recomendação ao Programa de Mestrado.

Ao professor Marcelino Andrade Gonçalves que oportunizou a realização de atividades do curso no campus da UFMS em Nova Andradina.

Aos colegas do Mestrado, em especial ao Douglas Nantes Gualberto, pelo apoio nos estudos e nas atividades desenvolvidas em conjunto, e ao Leandro de Jesus, que solícitamente compartilhou sua experiência e materiais para o desenvolvimento da minha pesquisa.

Aos professores do IFMS, que me auxiliaram com a distribuição da carga horária e organização dos horários ao longo deste período, em especial ao prof. Danilo Adriano Mickuci e Evandro Luis Souza Falleiros, que solícitamente sempre atenderam aos meus pedidos.

Aos meus amigos Anderson Lima e Willames Bezerra, que escancararam as portas das suas residências, me abrigando por diversas vezes, em especial na ocasião da minha mudança para Campo Grande.

Ao meu amigo Olavo José Luiz Jr., pela companhia nos estudos, nas viagens intermináveis no trecho Nova Andradina-Campo Grande e pela amizade incondicional.

Aos meus amigos do “DINF”, pelo apoio em todos os momentos desta jornada, antes mesmo dela começar, em especial ao Claudio Zarate Sanavria, que me incentivou e me colocou no caminho da pesquisa.

E a minha esposa, Fernanda Moretti, que mesmo abrindo mão da minha companhia, me apoiou durante todo este período.

*"A mão no pulso do tempo e o ouvido no coração de Deus."
(Pe. José Kentenich, sobre a aceitação dos desafios do mundo atual)*

Resumo

Este trabalho aborda a identificação automática do comportamento bovino, utilizando os dados de movimentação e posicionamento do animal e algoritmos de classificação supervisionada. Um colar composto por GPS, acelerômetro, giroscópio e magnetômetro foi desenvolvido e utilizado no pescoço de novilhas, mantidas no pasto experimental da Embrapa Gado de Corte, na cidade de Campo Grande / MS. Quatro unidades do colar coletaram mais de 500 horas de dados, sendo 71 horas com observação e anotação simultânea de um pesquisador. Os dados dos colares e das observações foram unificados e utilizados para construir o modelo de classificação automática dos comportamentos. Para classificar quatro comportamentos primário (Pastando/Procurando, Andando, Em Pé e Deitado), o algoritmo *Random Forest* obteve acurácia média de 93,99%. Separando os momentos que o animal estava Em Pé / Deitado Parado ou Ruminando, o modelo apresentou acurácia média de 89,55%. O trabalho também propõe uma metodologia para classificação dos comportamentos utilizando os dados dos colares e outras variáveis, tais como locais de interesse ou comportamentos já identificados. O comportamento Bebendo Água foi classificado utilizando as coordenadas geográficas e apresentou acurácia de 94,1%, enquanto sem elas a acurácia foi de apenas 71,2%. Esta pesquisa é fruto da parceria entre a FACOM / UFMS e a Embrapa Gado de Corte, dentro da linha de pesquisa de Tecnologias Computacionais para Pecuária.

Palavras-chaves: pecuária de precisão; comportamento bovino; reconhecimento de padrões; sensores; etologia.

Abstract

This work approaches the automatic identification of cattle behavior, using animals' movement and positioning data and supervised classification algorithms. A collar assembled with GPS, accelerometer, gyroscope and magnetometer was developed and it was used in the neck of heifers kept in the experimental pasture at Embrapa, in Campo Grande / MS. Four collars collected more than 500 hours of data, 71 hours with simultaneous viewing and annotation of a researcher. Data from collars and observations were unified and used to build a model for behavior automatic classification. To classify the four primary behaviors (Grazing / Looking, Walking, On Foot and laying), Random Forest algorithm had an average accuracy of 93.99%. By separating the moments the animal was On Foot or Laying in behaviors Stationary or Ruminating, the model had a mean accuracy of 89.55%. This work also proposes a methodology for classification of behaviors using data from collars and other variables, such as places of interest or behaviors already identified. The Drinking Water behavior was classified using geographic coordinates, and it presented an accuracy of 94.1%, while without the coordinates, the accuracy was only 71.2%. This research is a result of the partnership between FACOM / UFMS and Embrapa Beef Cattle, within Computational Technologies for Livestock research line.

Key-words: precision livestock farming; cattle behavior; pattern recognition; sensors; ethology.

Lista de figuras

Figura 1 – Hardware básico de um nó sensor (LOUREIRO <i>et al.</i> , 2003)	28
Figura 2 – Características do Acelerômetro, Giroscópio e Magnetômetro. Adaptado de STMicroelectronics (2013b) e STMicroelectronics (2013a).	31
Figura 3 – Processo de descoberta de conhecimento (TAN; STEINBACH; KUMAR, 2009).	39
Figura 4 – Abordagem geral para a construção de um modelo de classificação (TAN; STEINBACH; KUMAR, 2009).	40
Figura 5 – Exemplo de árvore de decisão: classificação de frutas baseado no sabor, cor, formato e tamanho (DUDA; HART; STORK, 2000).	41
Figura 6 – Exemplo de uma rede neural artificial multicamadas (TAN; STEINBACH; KUMAR, 2009).	42
Figura 7 – Visão lógica do método de aprendizagem de grupo (TAN; STEINBACH; KUMAR, 2009).	42
Figura 8 – Diagrama de fluxo de um processo de seleção de subconjuntos de características (TAN; STEINBACH; KUMAR, 2009).	44
Figura 9 – Tela inicial do WEKA.	48
Figura 10 – Tela de pré-processamento do WEKA.	48
Figura 11 – Tela com as opções do método de classificação do WEKA.	49
Figura 12 – Módulo GPS utilizado - Sparkfun Venus GPS. Disponível em Electronics (2015c).	57
Figura 13 – MiniIMU 9DOF: acelerômetro, giroscópio e magnetômetro. Disponível em Electronics (2015a).	58
Figura 14 – Resistor Variável de Luminosidade (LDR). Disponível em Webtronico (2015).	58
Figura 15 – Módulo OpenLog. Disponível em Electronics (2015b).	58
Figura 16 – Arduino Fio. Disponível em Arduino (2015).	59
Figura 17 – Projeto do circuito desenvolvido no <i>Fritzing</i>	59
Figura 18 – Quatro unidades da placa de circuito durante o processo de produção.	60
Figura 19 – Placa de circuito fixada na tampa da caixa.	61
Figura 20 – Cinto com a caixa do circuito.	62
Figura 21 – Sensor LDR e Antena GPS fixados ao colar.	62
Figura 22 – Circuito montado na protoboard para testes de bancada.	63
Figura 23 – Placa de circuito com os sensores encaixados.	63
Figura 24 – Área do Mangueiro Digital.	64
Figura 25 – Duas versões da caixa para encapsulamento do circuito.	65

Figura 26	– Posição ideal da antena GPS e LDR com relação a caixa com o circuito.	66
Figura 27	– Animal com o colar em posição invertida: amarrado no lado direito.	67
Figura 28	– Animal com o pescoço reto. Destaque para a posição de cada eixo do acelerômetro.	70
Figura 29	– Caixa com o circuito aberto, com destaque para a posição da bateria.	70
Figura 30	– Animal nas posições dos quatro comportamentos primários: (A) Pastando/Procurando, (B) Andando, (C) Em Pé e (D) Deitado.	71
Figura 31	– Troca de colar realizada no mangueiro com auxílio dos profissionais da Embrapa.	73
Figura 32	– Tela do Ápis na versão desenvolvida por Jesus (2014).	75
Figura 33	– Tela inicial para personalização da quantidade de animais e comportamentos observados.	76
Figura 34	– Tela com os botões para seleção dos animais.	77
Figura 35	– Tela com os botões para seleção dos comportamentos.	77
Figura 36	– Comportamentos primários identificados a partir dos dados de posição e movimentação.	85
Figura 37	– Proposta de modelo para classificação de comportamentos secundários.	94
Figura 38	– Esquema para classificação do comportamento Ruminando.	94
Figura 39	– Esquema para classificação do comportamento Bebendo Água.	99
Figura 40	– Plotagem da localização dos animais na área de pasto baseado nas coordenadas geográficas: em destaque os registros rotulados como Bebendo Água.	100

Lista de tabelas

Tabela 1	– Exemplo de matriz de confusão.	45
Tabela 2	– Interpretação dos valores do índice Kappa (LANDIS; KOCH, 1977).	47
Tabela 3	– Características dos animais utilizados no experimento.	69
Tabela 4	– Volume de Dados Coletados em Dias/Horas	73
Tabela 5	– Detalhes das coletas e observações organizados por colar.	74
Tabela 6	– Número de anotações realizadas por comportamento em cada colar.	79
Tabela 7	– Número de registros das observações com combinações nos colares.	79
Tabela 8	– Estrutura dos registros LAGM com a identificação do colar e valores dos sensores LDR e de movimentação.	80
Tabela 9	– Exemplo dos dados originais do colar: registros LAGM + GPS.	80
Tabela 10	– Exemplo de um registro do colar após tratamento.	80
Tabela 11	– Registros gerados pelo software de observação.	81
Tabela 12	– Registros do software de observação após tratamento.	81
Tabela 13	– Arquivo com os registros unificados e com o cabeçalho formatado para a ferramenta WEKA.	81
Tabela 14	– Relação dos 57 algoritmos utilizados para construção dos modelos de classificação do comportamento animal.	84
Tabela 15	– Detalhes da quantidade de comportamentos de cada um dos conjuntos de dados utilizados para construção e validação do modelo de classificação.	85
Tabela 16	– Algoritmos com melhor acurácia na validação utilizando o método <i>cross-validation</i> com 10 <i>folds</i> com todos os registros.	86
Tabela 17	– Detalhes do arquivo de teste com 15% dos dados de cada conjunto.	86
Tabela 18	– Resultados da validação individual dos conjuntos utilizando o método <i>Holdout</i>	87
Tabela 19	– Matriz de confusão do modelo gerado pelo conjunto B2 e o algoritmo <i>Random Forest</i> e validado com o método <i>Holdout</i>	87
Tabela 20	– Matriz de confusão do modelo gerado pelo conjunto C4 e o algoritmo <i>Random Forest</i> e validado com o método <i>Holdout</i>	87
Tabela 21	– Matriz de confusão com percentual de acerto da validação do modelo gerado com o conjunto A1 e o algoritmo <i>Random Forest</i>	88
Tabela 22	– Conjuntos agrupados por animal a partir dos dados dos conjuntos individuais.	89
Tabela 23	– Resultado da validação dos conjuntos agrupados por animal utilizando o método <i>Holdout</i>	90

Tabela 24	– Conjuntos agrupados por colar a partir dos dados dos conjuntos individuais.	90
Tabela 25	– Resultado da validação dos conjuntos agrupados por colar utilizando o método <i>Holdout</i>	91
Tabela 26	– Matriz de confusão da validação do modelo de classificação dos comportamentos primários utilizando o método <i>cross-validation</i>	92
Tabela 27	– Comparação da acurácia de cada classe nos modelos gerados com os algoritmos <i>Random Forest</i> e SVM.	93
Tabela 28	– Matriz de confusão do modelo de classificação do comportamento ruminando junto aos comportamentos primários.	95
Tabela 29	– Matriz de confusão do modelo de classificação do comportamento Em Pé-Ruminando.	96
Tabela 30	– Matriz de confusão do modelo de classificação do comportamento Deitado-Ruminando.	96
Tabela 31	– Acurácia da classificação dos comportamentos Em Pé-Ruminando e Em Pé-Parado.	97
Tabela 32	– Acurácia da classificação dos comportamentos Deitado-Ruminando e Deitado-Parado.	97
Tabela 33	– Matriz de confusão com acurácia do modelo de classificação do comportamento Bebendo Água como comportamento primário.	98
Tabela 34	– Matriz de confusão da classificação dos comportamentos primários e Bebendo Água, utilizando as coordenadas geográficas.	98
Tabela 35	– Matriz de confusão da classificação do comportamento Bebendo Água agregando os demais comportamentos em uma única classe.	99
Tabela 36	– Matriz de confusão da classificação do comportamento Bebendo Água apenas com os registros filtrados na área do bebedouro.	100
Tabela 37	– Comparativo da acurácia nas quatro estratégias utilizadas para classificação do comportamento Bebendo Água.	101
Tabela 38	– Comparativo dos resultados obtidos com os apresentados por Martiskainen <i>et al.</i> (2009).	101
Tabela 39	– Comparativo do resultado da validação do modelo de classificação dos comportamentos primários com os apresentados por Nadimi <i>et al.</i> (2012) e Jesus (2014).	102
Tabela 40	– Comparativo dos resultados obtidos com os apresentados por Dutta <i>et al.</i> (2015).	102
Tabela 41	– Comparativo dos resultados obtidos com os apresentados por González <i>et al.</i> (2015).	103

Lista de abreviaturas e siglas

API	<i>Application Programming Interface</i> - Interface de Programação de Aplicações
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária
FACOM	Faculdade de Computação
GNSS	<i>Global Navigation Satellite System</i> - Sistema Global de Navegação por Satélite
GPS	<i>Global Positioning System</i> - Sistema de Posição Global
GUI	<i>Graphical User Interface</i> - Interface Gráfica do Usuário
IBGE	Instituto Brasileiro de Geografia e Estatística
KDD	<i>Knowledge Discovery in Database</i> - Conhecimento Descoberto em Base de Dados
LDR	<i>Light Dependent Resistor</i> - Resistor Dependente de Luz
MANET	<i>Mobile Ad hoc Network</i> - Rede Móvel Ad hoc
MAPA	Ministério da Agricultura, Pecuária e Abastecimento
PIB	Produto Interno Bruto
RSSF	Redes de Sensores Sem Fio
UFMS	Universidade Federal de Mato Grosso do Sul
WEKA	<i>Waikato Environment for Knowledge Analysis</i>

Sumário

1	Introdução	21
1.1	Motivação	23
1.2	Objetivos	24
1.3	Organização do texto	24
2	Fundamentação Teórica	27
2.1	Redes de Sensores Sem Fio	27
2.1.1	Componentes das RSSFs	27
2.1.2	Aplicação das RSSFs	28
2.1.3	Características de RSSFs	29
2.1.4	Limitação da Energia Disponível	30
2.2	Sensores Utilizados	30
2.2.1	Sensores de Movimentação	31
2.2.2	Sensores de Luminosidade	32
2.3	Comportamento Animal	33
2.3.1	Identificação do Comportamento Animal	33
2.3.2	Principais Comportamentos do Gado	34
2.3.3	Ferramentas para Identificação do Comportamento Animal	35
2.3.4	Instrumentos para Coleta de Dados	36
2.3.5	Influência do Ambiente no Comportamento Animal	37
2.4	Reconhecimento de Padrões	37
2.4.1	Mineração de Dados	38
2.4.2	Modelos de Classificação	39
2.4.3	Técnicas para a Construção de Modelos de Classificação	40
2.4.4	Overfitting de Modelo	43
2.4.5	Seleção de Atributos (Dimensionalidade)	43
2.4.6	Metodologia para Avaliação de Modelos	44
2.4.7	Métricas para Avaliação de Modelos	45
2.4.8	Ferramenta WEKA	47
2.5	Considerações Finais	49
3	Trabalhos Relacionados	51
3.1	Equipamentos e Sensores Utilizados	51
3.2	Coleta de Dados	52
3.3	Comportamentos Observados	52
3.4	Tratamento e Processamento dos Dados	54

3.5	Resultados	55
3.6	Considerações Finais	56
4	Produção dos Colares	57
4.1	Sensores Utilizados	57
4.2	Confecção da Placa de Circuito	59
4.3	Confecção dos Colares	61
4.4	Implementação e Testes	62
4.4.1	Circuito e Sensores	62
4.4.2	Bateria	63
4.5	Indicadores da Experiência em Campo	64
4.5.1	Material Utilizado	65
4.5.2	Posição do Colar no Animal	66
4.6	Considerações Finais	67
5	Experimento em Campo	69
5.1	Animais Utilizados no Experimento	69
5.2	Posição do Colar no Animal	69
5.3	Comportamentos Observados	70
5.4	Período de Observação	72
5.5	Software para Anotação dos Comportamentos	75
5.6	Combinação dos Dados Coletados	78
5.7	Tratamento dos Dados Coletados	79
5.7.1	Dados do Colar	79
5.7.2	Dados do Software de Observação	80
5.7.3	Unificação dos Registros	81
5.8	Considerações Finais	82
6	Classificação dos Comportamentos	83
6.1	Processamento e Análise dos Dados	83
6.1.1	Ferramentas Utilizadas	83
6.1.2	Análise e Seleção dos Algoritmos	84
6.1.3	Avaliação dos Conjuntos de Dados Individuais	86
6.1.4	Avaliação de um Conjunto Descartado	88
6.1.5	Combinação dos Conjuntos de Dados	88
6.1.5.1	Variação dos Dados por Animal	89
6.1.5.2	Variação dos Dados por Colar	89
6.1.5.3	Análise dos Resultados das Variações	91
6.2	Modelos de Classificação dos Comportamentos Primários	91
6.2.1	Validação do Modelo Utilizando o Algoritmo Random Forest	92

6.2.2	Validação do Modelo Utilizando o Algoritmo SVM	92
6.3	Modelos de Classificação dos Comportamentos Secundários	93
6.3.1	Comportamento Ruminando	93
6.3.1.1	Classificação do Comportamento Ruminando como Com- portamento Primário	95
6.3.1.2	Classificação do Comportamento Em Pé-Ruminando	95
6.3.1.3	Classificação do Comportamento Deitado-Ruminando	96
6.3.1.4	Resultados da Classificação do Comportamento Ruminando	96
6.3.2	Comportamento Bebendo Água	97
6.3.2.1	Classificação do Comportamento Bebendo Água como Com- portamento Primário	97
6.3.2.2	Classificação do Comportamento Bebendo Água como Com- portamento Secundário	99
6.3.2.3	Resultados da Classificação do Comportamento Bebendo Água	100
6.4	Comparação dos Resultados com Outros Trabalhos	101
6.5	Considerações Finais	103
7	Conclusão	105
7.1	Trabalhos Futuros	106
	Referências	109

1 Introdução

Segundo o Ministério da Agricultura, Pecuária e Abastecimento (MAPA), o agromercado representa entre 22% e 23% do PIB (Produto Interno Bruto) total da economia brasileira, sendo a pecuária responsável por cerca de 30% do valor produzido (BRASIL, 2014). O rebanho bovino brasileiro, de acordo com o IBGE (Instituto Brasileiro de Geografia e Estatística), passa de 211 milhões de cabeças, com 33,6% do total localizado na região centro-oeste do país, sendo o estado de Mato Grosso do Sul o 4º maior rebanho do país, com mais de 21 milhões de animais (IBGE, 2013) e o 2º maior estado no número de animais abatidos (IBGE, 2015).

Até 2024, há uma projeção que prevê um aumento no consumo e na produção de carne bovina em quase 12 milhões de toneladas (MAPA, 2014). O aumento da produção gerará competição por mais terras, recursos naturais e mão de obra, sendo necessário o aprimoramento das técnicas, para aumentar e melhorar a produção. O segmento brasileiro de pecuária cresceu nos últimos 15 anos à taxa de 5% ao ano, porém, o atual ganho médio de peso do rebanho bovino está aquém das potencialidades do setor (MAPA, 2013). De acordo com o MAPA (2013), melhoramentos nos ganhos de produtividade, decorrentes de avanços nas áreas de melhoramento genético, controle de doenças, pastagens melhoradas e principalmente no sistema integração lavoura-pastagem ou lavoura-pastagem-floresta, permitirão que o país triplique a produtividade do rebanho bovino. A adoção dos princípios das práticas de manejo racional e de bem-estar, além de melhorar os atributos da carne e leite, proporciona maior eficiência econômica da propriedade, na facilitação do manejo e em maior rendimento de carcaça.

As mudanças na economia mundial também refletem sobre a pecuária de corte, exigindo que se aumente a produção e a produtividade da carne de qualidade, de maneira competitiva, eficiente e eficaz. O reflexo desta pressão é a busca por genótipos que se adaptem às condições climáticas do Brasil e com características produtivas semelhantes aos dos animais europeus (FILHO; FIGUEIREDO, 2003).

As principais subespécies bovinas são o *Bos taurus* (gado europeu) e *Bos indicus* (gado zebuino ou indiano). Na pecuária de corte nacional, predomina-se a criação do gado zebuino, sendo a raça Nelore responsável por 80% do rebanho (SILVEIRA, 2012). Porém, a carne desta espécie é reconhecidamente mais dura que a do gado europeu, fazendo com que a carne oriunda da espécie européia seja comercializada com preço superior.

Aumentar a produção de carne da espécie *Bos taurus* significa gerar maior receita com a sua comercialização. Sendo a pecuária de grande importância para a economia nacional e a região centro-oeste a maior produção do país, é importante avaliar a viabi-

lidade de criação da espécie no centro do Brasil. Mas um dos principais problemas para o pecuarista brasileiro é o carrapato. O parasita inocula substâncias no gado, pela saliva, causando coceira e diminuição do apetite e, com exceção dos estados da Região Sul, o problema ocorre durante o ano todo (FURLONG, 2005). O gado da espécie indiana é mais resistente aos carrapatos comparado ao gado europeu. Segundo Francis e Little (1964 apud ANDREOTTI, 2002), os bovinos europeus apresentam em média 10,5 vezes mais carrapatos que os de cruzamentos com zebuínos.

Desconsiderando o problema com os carrapatos, trazer a espécie *Bos taurus* para criação na região centro-oeste do Brasil traz junto algumas incertezas, pois não se sabe qual é o comportamento da espécie no cerrado brasileiro. Este cenário exige mudanças na produção e na gestão dos negócios agropecuários. A Pecuária de Precisão pode ser utilizada para auxiliar o pecuarista, a partir de um sistema de monitoramento que analise o comportamento animal, a fim de avaliar o bem-estar e a produção.

A Pecuária de Precisão é definida (CÁCERES *et al.*, 2011) como a gestão de bovinos a partir do uso da Tecnologia da Informação e Comunicação, a fim de garantir boas práticas na produção de carne. Os dados do rebanho, das áreas de pasto georreferenciadas e da automação da indústria da carne, possibilitam otimizar os custos de produção, reduzir os impactos ambientais e aumentar a produtividade. Segundo a Embrapa Gado de Leite (2012), a Pecuária de Precisão é uma quebra de paradigmas, uma vez que os dados passam a ser tratados individualmente para cada animal na tomada de decisão, diferente do modelo que considera apenas um valor de amostragem do rebanho. Assim, é possível monitorar o potencial genético de cada animal, baseado em metas econômicas ou em índices de bem-estar, levando a melhores estratégias de manejo.

A Embrapa (Empresa Brasileira de Pesquisa Agropecuária) é uma empresa vinculada ao MAPA, criada em 1973, com a missão de viabilizar soluções de pesquisa, desenvolvimento e inovação para a sustentabilidade da agricultura, em benefício da sociedade brasileira (EMBRAPA, 2015a). A unidade da Embrapa de pesquisa de Gado de Corte, localizada na cidade de Campo Grande, foi oficialmente inaugurada em 1977, no mesmo ano em que a cidade se tornou a capital do estado de Mato Grosso do Sul. A escolha da cidade foi defendida por especialistas do ramo por possuir condições favoráveis à produção de bovinos de corte (EMBRAPA, 2015b). A missão da unidade é viabilizar soluções tecnológicas sustentáveis para a cadeia produtiva da pecuária de corte, em benefício da sociedade brasileira (EMBRAPA, 2015c).

A Embrapa Gado de Corte e a FACOM/UFMS (Faculdade de Computação / Universidade Federal de Mato Grosso do Sul) possuem um histórico de parceria visando o desenvolvimento de tecnologias para Pecuária de Precisão. Fruto da parceria, diferentes ferramentas já foram desenvolvidas, tais como: o portal e-SAPI bovis, o sistema Taurus, a plataforma Thermus e o *middleware* Cerberus. Destaca-se dois trabalhos com foco na

identificação do comportamento animal: Oliveira (2013b) utilizou um conjunto de dados georreferenciados e um colar GPS (*Global Positioning System* - Sistema de Posição Global) para registrar o deslocamento do animal dentro da área de pasto e Jesus (2014) desenvolveu um colar de baixo custo, para identificação do comportamento do gado a partir dos dados de GPS.

A partir desta parceria, este trabalho foi proposto para desenvolver e aplicar tecnologias que auxiliem na resolução de problemas na área de Pecuária de Precisão, a partir da identificação automática dos comportamentos bovino.

1.1 Motivação

O gerenciamento das atividades na produção pecuária pode ser realizado de forma manual ou automática. As práticas manuais requerem detecção visual do animal, podem ter custo elevado e o tempo de retorno e qualidade das informações podem não ser aceitáveis, em função de possíveis erros gerados nos registros. Já as práticas automatizadas, possuem maior acurácia nos dados (redução nos erros de leitura), identificam a origem e o controle do número de animais, permitindo traçar rapidamente o histórico individual dos animais (RUIZ-GARCIA; LUNADEI, 2011).

Ruiz-Garcia e Lunadei (2011) destacam que a produção animal tem evoluído nos últimos anos em função do uso de ferramentas de precisão, utilizadas para prever eventos futuros, realizar diagnóstico preventivo e auxiliar na tomada de decisões. Os autores reforçam que apenas a automatização das atividades não é suficiente, já que o conhecimento agrônômico é essencial para o desenvolvimento da área. Assim, combinar estes conhecimentos com *softwares* de gerenciamento, que possibilitem o detalhamento das informações coletadas e as integre com ferramentas de gestão, é fundamental.

Apesar das vantagens, os autores destacam alguns desafios encontrados para o desenvolvimento de aplicações computacionais para uso no campo, em especial aquelas que utilizam sensoriamento. Entre eles, pode-se citar a condição agressiva do ambiente (exposição à excesso de sujeira e variações na umidade e temperatura); o orçamento de energia; a necessidade de detecção de falhas e isolamento de medições erradas; o nível de granularidade, baseado no detalhamento e acurácia da informação esperada; a falta de pessoal qualificado para atender a demanda; e a forte resistência dos produtores em compartilhar as informações.

Diferentes trabalhos ((SCHEIBE; GROMANN, 2006), (MARTISKAINEN *et al.*, 2009), (NADIMI *et al.*, 2012), dentre outros) abordaram a identificação automática do comportamento animal. Entretanto, não conseguiram identificar satisfatoriamente todos os comportamentos avaliados. Identificar um número maior de comportamentos é uma possibilidade em aberto, considerando que os trabalhos que buscaram classificar mais

comportamentos não alcançaram resultados satisfatórios.

Este trabalho utilizou sensores de movimentação e de posicionamento, a fim de prover funcionalidades que auxiliem a produção e a gestão pecuária, a partir da identificação automática dos comportamentos bovino.

1.2 Objetivos

O objetivo principal do trabalho é evoluir um sistema de identificação do comportamento bovino, incluindo novos sensores, para melhorar a acurácia da classificação dos comportamentos: Andando, Pastando/Procurando, Deitado e Em Pé, e classificar automaticamente outros comportamentos tais como Ruminando e Bebendo Água.

De maneira detalhada, os objetivos específicos são:

- Evoluir uma solução de *hardware* para coleta de dados do comportamento do gado, instalando no equipamento sensores de movimentação (acelerômetro, giroscópio e magnetômetro);
- Configurar e testar os novos sensores, identificando a posição ideal para instalá-lo, de maneira que os valores dos eixos fiquem padronizados;
- Elaborar um sistema para filtrar e tratar os dados coletados pelo equipamento;
- Avaliar diferentes algoritmos de classificação para identificar aqueles com melhor acurácia na classificação dos comportamentos Andando, Pastando/Procurando, Deitado e Em Pé;
- Identificar comportamentos secundários, a partir da combinação dos dados dos sensores com outras variáveis, por exemplo com locais de interesse na área de pasto;
- Realizar experimentos em campo, coletando os dados para classificar e validar a identificação automática dos comportamentos.

1.3 Organização do texto

O trabalho está organizado da seguinte forma: no Capítulo 2 são abordados os principais fundamentos envolvidos na proposta, apresentando-se as Redes de Sensores Sem Fio, as características dos sensores utilizados, os principais conceitos para trabalhar com identificação do comportamento animal e os conceitos aplicados da área de Reconhecimento de Padrões. No Capítulo 3, são apresentados os detalhes dos trabalhos utilizados como base para o desenvolvimento desta pesquisa. No Capítulo 4, os sensores utilizados e

o processo de confecção e testes dos colares são apresentados. No Capítulo 5, apresenta-se detalhes do trabalho em campo, como as informações dos animais e área utilizada, o processo de programação do *software* utilizado e do processo de tratamento dos dados coletados. O Capítulo 6 traz informações sobre o processamento e análise dos dados, bem como os modelos de classificação automática dos comportamentos propostos. Por fim, as conclusões sobre o trabalho desenvolvido e a proposta para trabalhos futuros são apresentadas.

2 Fundamentação Teórica

Para desenvolver este trabalho foram revisados os principais conceitos envolvidos, para identificar as possibilidades e limitações das tecnologias disponíveis. A Seção 2.1 descreve os principais fundamentos relacionados às redes de sensores. Na Seção 2.2, são apresentadas as características dos sensores de movimentação, de localização e de luminosidade, que foram utilizados neste trabalho. Na Seção 2.3, são apresentadas as características para o monitoramento do comportamento animal. Na Seção 2.4, é apresentada a área de Reconhecimento de Padrões, com os conceitos que foram aplicados para a classificação do comportamento bovino.

2.1 Redes de Sensores Sem Fio

Os avanços tecnológicos nos sistemas eletro-mecânicos, na comunicação sem fio e no desenvolvimento de sensores, permitiram a evolução de um tipo especial de rede sem fio, chamada de RSSF (Redes de Sensores Sem Fio).

Uma RSSF pode ser vista como um tipo especial de MANET (*Mobile Ad hoc Network* - Rede Móvel *Ad hoc*), em que os elementos da rede se comunicam diretamente por meio de enlace de comunicação sem fio. A diferença entre elas está relacionada às tarefas que cada nó executa. Enquanto na MANET os nós executam tarefas distintas, na RSSF eles tendem a executar uma tarefa conjunta, baseado no esforço colaborativo de um grande número de nós (AKYILDIZ *et al.*, 2002).

2.1.1 Componentes das RSSFs

As RSSFs são compostas basicamente pelos nós sensores, as interfaces de comunicação sem fio e os nós *gateway*. Cada um dos nós sensores possuem capacidade de sensoriamento multifuncional, processamento e comunicação. Os componentes básicos dele são: sensor, processador, memória, rádio e bateria, conforme mostra a Figura 1.

Uma RSSF pode ser formada por diferentes sensores. Cada sensor apresenta características físicas e de funcionamento próprias. Loureiro *et al.* (2003) indicam que muitos modelos de dispositivos sensores compartilham duas características:

- (i) habilidade de sensoriamento diminui quando a distância aumenta; e
- (ii) a habilidade de sensoriamento pode melhorar com o tempo de exposição, considerando os efeitos decrescentes dos ruídos nas medições.

Para os sensores que são posicionados longe do fenômeno analisado, são necessárias

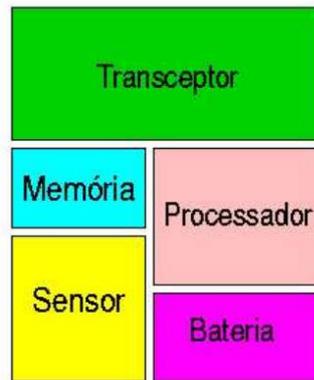


Figura 1 – Hardware básico de um nó sensor (LOUREIRO *et al.*, 2003)

técnicas complexas para filtrar os ruídos. Em geral, os nós sensores são espalhados e se comunicam com um nó central, que armazena e processa os dados (AKYILDIZ *et al.*, 2002). A posição de cada nó não precisa ser planejada ou predeterminada, já que uma das características das RSSF é de se auto-organizar. Quando um sensor é adicionado ou removido, o próprio sistema deve reorganizar a sua topologia de comunicação.

Em alguns casos, a RSSF pode ter nós com a função de atuadores, que possuem a capacidade de alterar valores do ambiente, para controlar o objeto monitorado. Quando um nó possui a função de monitoramento e atuação, ele é chamado de transdutor (LOUREIRO *et al.*, 2003).

Os nós sensores podem processar localmente os dados coletados e então enviá-los a uma estação base, através de um nó *gateway*. O nó *gateway* tem a função de permitir a comunicação da rede de sensores com outras redes, para que os dados cheguem até o computador que roda a aplicação principal.

Os nós sensores, além de coletar dados do ambiente, podem atuar como roteadores dos dados gerados por outros nós da rede. Nestes casos os nós são chamados de *sink node* ou nós sorvedouros.

2.1.2 Aplicação das RSSFs

Os nós de uma RSSF podem ser compostos por um ou mais sensores e coletar diferentes dados de um ambiente. Eles devem ser pequenos e com baixo consumo de energia, permitindo sua aplicação em diferentes áreas.

Akyildiz *et al.* (2002) citam algumas das áreas de possível aplicação das RSSFs: aplicações militares (monitoramento de equipamentos e munições; vigilância do campo de batalha; reconhecimento de forças e terrenos opostos; avaliação de danos de batalha; detecção de ataques químicos e biológicos); aplicações ambientais (rastreamento de animais; monitoramento das condições ambientais que afetam a agricultura e a pecuária; irrigação;

agricultura de precisão); aplicações de saúde (rastreamento e monitoramento dos pacientes dentro de um hospital; administração de medicamentos em hospitais; telemonitorização de dados fisiológicos humanos); aplicações domésticas (ambientes adaptativos ao usuário).

Dentre os diferentes cenários, destaca-se a aplicação no monitoramento de ambientes, que segundo [Loureiro *et al.* \(2003\)](#) pode ser aplicado tanto em ambientes internos quanto externos, a fim de determinar os valores da temperatura, pressão atmosférica, quantidade de luz e umidade.

2.1.3 Características de RSSFs

As principais características de uma RSSF, comparada às redes de computadores tradicionais e as redes *ad hoc*, são: possuir um grande número de nós distribuídos; ter restrições de energia; ter limitações em potência, capacidade computacional e de memória nos nós e ser propensos a falhas; alterar frequentemente a topologia da rede; possuir mecanismos para auto-configuração e adaptação devido a problemas como falhas de comunicação e perda de nós.

Essas características particulares das RSSFs variam em função da aplicação a qual a rede está atendendo. Algumas delas são apresentadas a seguir, baseadas em [Loureiro *et al.* \(2003\)](#) e [Akyildiz *et al.* \(2002\)](#):

Endereçamento dos nós sensores: algumas aplicações exigem a identificação de cada nó individualmente, enquanto para outras apenas o resultado coletado interessa, independente da sua identificação.

Agregação dos dados: dados coletados são sumarizados para posteriormente serem enviados ao nó central, reduzindo o número de mensagens transmitidas pela rede.

Mobilidade dos sensores: nós podem ser fixos ou móveis dentro do ambiente. nós para coletar a umidade ou temperatura de um ambiente podem ser fixos, enquanto para monitorar a posição de um animal dentro de uma área, devem ser móveis.

Quantidade de sensores: as aplicações podem utilizar de dezenas a milhares de sensores. Quanto maior a necessidade e interesse na escalabilidade do sistema, maior o número de sensores.

Custos de produção: o custo global da rede está relacionado ao número e o custo dos nós existentes. O custo total não deve superar o custo da implantação de sensores tradicionais para que ela seja viável financeiramente.

Restrições de hardware: um nó sensor é constituído por quatro componentes básicos: unidade de sensoriamento, unidade de processamento, unidade de comunicação e unidade de energia. Dependendo da aplicação podem existir componentes adicionais. Mesmo assim é fundamental que o sistema continue com baixo consumo de energia, baixo

custo de produção, seja autônomo e adapte-se ao ambiente.

Tolerância a falhas: os nós podem falhar, por danos físicos ou interferência do ambiente, ou ser bloqueados por falta de energia. A falha de um nó pode afetar a tarefa global do sistema. A rede deve possuir a capacidade de manter as suas funcionalidades sem qualquer interrupção quando algum nó falhar.

Limitação da energia disponível: em grande parte das aplicações os sensores são colocados em áreas remotas, que não estão facilmente acessíveis para manutenção. O tempo de vida do nó é diretamente relacionado à quantidade de energia disponível.

2.1.4 Limitação da Energia Disponível

[Akyildiz et al. \(2002\)](#) dividem o consumo de energia da rede em três domínios, baseado nas principais funções dos nós sensores: sensoriamento, comunicação e processamento de dados.

As atividades de sensoriamento estão relacionadas com a percepção do ambiente e a coleta de dados. O consumo de energia varia em função da distância do alvo, ruídos do ambiente, tipo de dado coletado, volume de informação e frequência de amostragem ([LOUREIRO et al., 2003](#)). Estas variações estão relacionadas ao tipo de aplicação: algumas necessitam de monitoramento constante dos eventos, enquanto outras não.

Os autores dividem o processamento na rede de sensores em duas categorias: processamento de suporte e processamento da informação. O primeiro está relacionado ao gerenciamento, comunicação e manutenção da rede. O segundo com a compressão, correlação e criptografia dos dados, por exemplo. Os dados podem ou não ser processados pelo nó sensor, baseado na função da aplicação ou no envolvimento do nó com a rede.

O maior custo energético da rede está no domínio da comunicação. A energia consumida para comunicação, entre dois pontos, está diretamente relacionada à distância entre eles. O processamento de dados, antes da comunicação, é importante para reduzir o tamanho dos pacotes a serem transmitidos. O gasto energético em processamento de dados é muito menor do que com a comunicação de dados ([AKYILDIZ et al., 2002](#)).

2.2 Sensores Utilizados

Os trabalhos apontam o GPS e o acelerômetro, como os principais sensores utilizados para aferir o comportamento animal. Neste trabalho, além deles, utiliza-se outros três sensores: giroscópio, magnetômetro e de luminosidade. As características destes sensores são descritas nesta seção.

2.2.1 Sensores de Movimentação

O acelerômetro e o giroscópio são sensores de aceleração, com a capacidade de medir a aceleração linear e a aceleração angular, respectivamente, em um, dois ou três eixos dos objetos ao qual está acoplado. Estes tipos de sensores têm sido utilizados há décadas em aeronaves, navios, veículos e robôs. Com a redução do tamanho, peso e custo, novas possibilidades de aplicações surgiram, entre elas o monitoramento da atividade humana (ALTUN; BARSHAN; TUNÇEL, 2010) e da atividade animal. O magnetômetro pode ser usado para determinar o ângulo no qual o dispositivo é girado, relativo ao polo magnético norte da Terra.

Juntando e processando as medidas dos três sensores é possível construir um medidor completo de posição, velocidade e aceleração. No caso de animais livres na natureza, a medição simultânea de diferentes parâmetros pode ajudar a melhorar a identificação de diferentes comportamentos (WATANABE *et al.*, 2005). O sistema de medição de aceleração pode ser aplicado para investigar reações emocionais ou movimentos complexos de animais; também pode ser utilizado em paralelo com outros registros, para indicar perturbações externas ou nervosismo (SCHEIBE; GROMANN, 2006).

O acelerômetro é um sensor que registra o quão rapidamente a velocidade de um corpo varia ao longo do tempo. A partir dele pode-se determinar movimentos nos eixos X (para frente e para trás), Y (para direita e para esquerda) e Z (para cima e para baixo). A unidade de medida dele é baseado na força g , onde $1g$ equivale a $9,8 \text{ m/s}^2$. O giroscópio é um dos componentes do sistema de navegação inercial e consiste em servir como referência de direção, uma vez que opõe-se a qualquer tentativa de mudar a sua direção original. Ele indica a variação da inclinação do movimento dentro de cada eixo, utilizando graus/segundo como a unidade de medida. Já o magnetômetro é um instrumento que mede os campos magnéticos e determina a força, orientação e direção de um campo magnético, que utiliza a unidade *Gauss*. A Figura 2 apresenta as características dos sensores.

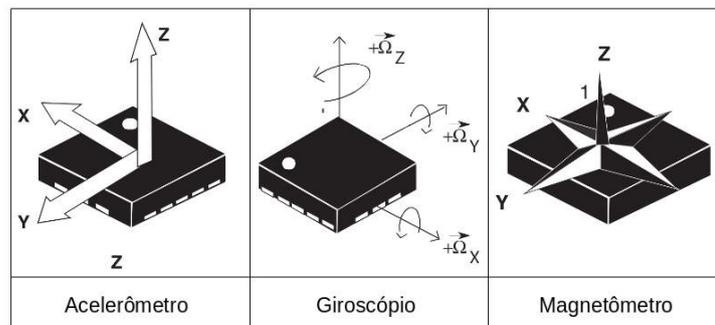


Figura 2 – Características do Acelerômetro, Giroscópio e Magnetômetro. Adaptado de [STMicroelectronics \(2013b\)](#) e [STMicroelectronics \(2013a\)](#).

Para [Scheibe e Gromann \(2006\)](#) sensores com medição tridimensional são sufici-

entamente sensíveis e convenientes para registrar os movimentos de gado livre, enquanto sistemas com medições de aceleração bidimensionais não são suficientes para uma descrição completa. Os autores indicam que um sensor com sensibilidade de $4g$ é suficiente para registrar os dados do movimento do gado, considerando que em seu trabalho poucos registros geraram *overflow* (o valor do movimento ultrapassou o valor máximo de sensibilidade do sensor) e que a maior parte dos registros tinha força inferior a $2g$. Os dados eram gravados em um memória no próprio dispositivo, em uma taxa de amostragem de 10 ms, em arquivos de texto, contendo o horário da gravação e os três valores de aceleração registrados. O armazenamento local dos dados permite que dados de diferentes sensores sejam gravados e encaminhados por uma rede sem fio ou descarregados diretamente em um computador.

Em [Watanabe et al. \(2005\)](#) os sinais dos eixos do acelerômetro foram zerados quando o animal estava deitado na posição horizontal. Em todos os ensaios realizados repetiu-se o procedimento. O trabalho avaliou os comportamentos de um gato doméstico, sendo que este processo de calibração foi facilitado, considerando o porte do animal. Repetir o procedimento em animais de maior porte, como um boi, é um desafio.

2.2.2 Sensores de Luminosidade

Os sensores de luminosidade são baseados em um resistor variável dependente de luz (LDR - *Light Dependent Resistor*), cuja resistência varia linearmente com a intensidade de luz incidente. Um LDR converte a luz em valores de resistência, de maneira que sua resistência diminui quando a luz é muito alta e aumenta quando a intensidade é baixa.

A partir dele pode se indicar o índice de luz, diferenciando áreas de sombra e de sol, baseado na quantidade de luz irradiada no ambiente. A unidade utilizada para medição de intensidade luminosa é o *lux*. Em geral, zero *lux* representa ausência de luz e indica uma resistência entre entre 1 e $20M\Omega$, enquanto 100 *lux*, que indica uma área com luminosidade, registra cerca de $5K\Omega$.

Os sensores LDR não são precisos, variando os valores de leitura de uma unidade para outra. Porém, existem sensores digitais que podem ser utilizados em diferentes aplicações, com maior acurácia e que podem ser configurados com diferentes níveis de detecção, com limites entre 0 e 40.000 *lux*.

O TSL2560 e TSL2561 são exemplos de sensores de conversão de iluminação digital. Esse tipo de sensor transforma a intensidade de luz para uma saída de sinal digital com interface I2C ou SMBus. O dispositivo combina um fotodiodo de banda larga e um fotodiodo infravermelho em um único circuito integrado, que por meio de um conversor Analógico-Digital transforma as correntes do fotodíodo em um valor que representa a irradiância medida. O valor pode ser convertido para *lux* e utilizado para aplicações de

uso geral com interesse na detecção de luz (TAOS, 2005).

2.3 Comportamento Animal

A etologia é a parte da ecologia que trata dos hábitos dos animais e da acomodação dos seres vivos às condições do ambiente, segundo o dicionário Michaelis¹. Chama-se de comportamento aquilo que percebe-se das reações de um animal ao ambiente que o cerca, segundo Carthy (1969 apud FERRAZ, 2011).

Para o estudo do comportamento é necessário observações dos movimentos e postura do animal. Além disso, é preciso obter dados precisos sobre suas ações, a fim de compreender em que condições elas acontecem. A identificação de comportamentos pode servir para o gerenciamento dos animais e das áreas em que eles vivem. Scheibe e Gromann (2006) indicam que desvios nos padrões normais de comportamento do animal, podem ser resultados de uma doença ou de perturbações exteriores.

2.3.1 Identificação do Comportamento Animal

O problema inicial para qualquer tipo de trabalho, que busque a identificação do comportamento, é estabelecer quais os comportamentos habituais da espécie. Uma abordagem é estudar as populações de animais selvagens em um espaço sem a interferência dos seres humanos. Porém, o gado domesticado no pasto possui como ancestral o *aurochs*, uma espécie que desapareceu em 1627 segundo Clutton-Brock (1999 apud KILGOUR *et al.*, 2012), impossibilitando que esta seja observada. A fim de orientar as condições a serem consideradas para a observação dos comportamentos, a *British Farm Animal Welfare Council* (KILGOUR, 2012), elaborou uma lista que indica os melhores padrões de bem-estar do gado, considerados normais para a sua criação, para serem utilizados na avaliação e estudo do bem-estar animal.

Para formar uma lista de comportamentos do gado (etograma) baseado nos padrões de bem-estar, Kilgour (2012) analisou 22 trabalhos sobre o bem-estar animal, realizados entre 1927 e 2009, e apresentou 40 comportamentos identificáveis do gado, classificados em cinco categorias: comportamentos de manutenção (p. ex. pastando ou ruminando), auto-expressão (p. ex. raspando a pata no chão ou esfregando o pescoço em algum objeto), social (p. ex. lambendo outro animal ou amamentando), procriação (p. ex. inseminando ou parindo) e social mútuo (p. ex. parado ou andando ao lado de outro animal). Apesar da diversidade, os comportamentos mais comuns observados foram: pastando, parado, ruminando, andando, bebendo água e lambendo algum tipo de suplemento. Os três primeiros ocupam entre 90% e 95% do dia do animal.

¹ Michaelis Moderno Dicionário da Língua Portuguesa. Disponível em: <http://michaelis.uol.com.br>

2.3.2 Principais Comportamentos do Gado

A maioria dos trabalhos revisados por Kilgour (2012), considera apenas os comportamentos pastando, parado e ruminando, ignorando outros comportamentos do animal. Apenas três dos 22 estudos revisados tinham informações detalhadas sobre os diversos comportamentos: Herbel e Nelson (1966), Zemo e Klemmedson (1970) e Hall (1989). Apesar dos dados sobre outros comportamentos do gado serem escassos na literatura, a maior parte dela indica que pastando é o comportamento mais comum.

Além das limitações dos trabalhos, quanto à restrição de comportamentos avaliados, outros problemas encontrados são: o pequeno tempo de observação, número baixo de animais no rebanho, áreas de pasto restritas e observação de um único animal do rebanho. Kilgour *et al.* (2012) realizou a observação do comportamento utilizando seis rebanhos em cinco propriedades diferentes, em que todos os animais foram observados, em uma área de pastagem maior que 200 hectares, durante seis sessões nos períodos da manhã e da tarde, realizado em dois meses (agosto e setembro) de dois anos consecutivos.

No trabalho de Kilgour *et al.* (2012) foi utilizado o conceito de comportamentos primários e secundários, em que um comportamento secundário é aquele que acontece ao mesmo tempo que um comportamento primário. Os comportamentos primários, definidos pelos autores, foram: Em Pé e Deitado. Ao todo foram observados 18 comportamentos. Em nenhum momento, em qualquer um dos seis rebanhos, todos os animais estavam no mesmo comportamento simultaneamente. Isso reforça o fato de que é necessário considerar e compreender a variação de comportamento de um animal para outro. Alguns trabalhos, avaliados por Kilgour *et al.* (2012), indicaram momentos em que todos os animais estavam pastando simultaneamente, porém em todos eles poucos animais foram observados. Hancock (1954 apud ANDERSON; CIBILS, 2013) indica que mesmo gados de um mesmo rebanho e de uma mesma cria, apresentam diferenças comportamentais.

Considerando o alto custo para equipar todos os animais de um rebanho, Oudshorn, Kristensen e Nadimi (2008 apud ANDERSON; CIBILS, 2013) apontam que é possível generalizar o comportamento de uma amostragem para todo o rebanho, porém não há estudos que indiquem o número exato de animais de um grupo a ser observado, para descrever e generalizar com acurácia o comportamento do grupo investigado. Para Otis e White (1999 apud ANDERSON; CIBILS, 2013), uma amostra de seis a doze animais pode ser considerada adequada, se planejada com base em evidências de correlação.

Apesar do trabalho de Kilgour *et al.* (2012) eliminar alguns dos problemas encontrados em trabalhos anteriores, a pesquisa não considerou o comportamento do animal no período noturno e em condições de chuva. Além disso, a atividade de ruminção foi considerada juntamente com a atividade de descanso (parado), já que os observadores não conseguiam ver a boca do animal (observações visuais realizadas sem uso de tecnologia

avançada). Mesmo com estas restrições, confirmou-se os resultados de outros trabalhos, quanto às principais atividades do gado: a atividade de pastagem representou 51% do tempo (6,1h das 12h observadas), seguido pelas atividades de descanso (em pé ou deitado) e andando. Dentre os seis rebanhos avaliados, a variação da proporção total destas atividades foi pequena (entre 95,4% e 97,7% do tempo total observado).

Nos trabalhos avaliados por [Kilgour \(2012\)](#), que registraram e separaram os comportamentos ocorridos durante o dia e a noite (11 dos 22 trabalhos), a atividade pastando aconteceu com maior frequência durante a luz do dia, com poucos registros durante a noite. O pico da atividade pastando se dá no nascer e pôr do sol, informação confirmada em cinco dos seis rebanhos avaliados por [Kilgour et al. \(2012\)](#). Já a atividade de ruminação, registrada tanto quando o animal estava em pé quanto deitado, ocorre com maior frequência no período noturno, assim como a atividade de descanso.

2.3.3 Ferramentas para Identificação do Comportamento Animal

Os primeiros estudos sobre comportamento animal (em torno de 1920) foram realizados utilizando a observação humana e registros manuscritos das atividades. Com a evolução do homem e o surgimento de novas tecnologias, novas abordagens foram adotadas, sendo que atualmente o uso de câmeras de vídeo e binóculos, com tecnologia de visão noturna, tem sido utilizadas como ferramentas de apoio ([ANDERSON; CIBILS, 2013](#)).

Apesar destas serem ferramentas eficientes, existem problemas relacionados as limitações do homem, como a fadiga em trabalhos que exigem longos períodos de observação e a influência da presença do observador no habitat do animal. O uso de equipamentos eletrônicos, além de substituir a presença do homem, pode melhorar a qualidade dos dados coletados.

Há registros que já em 1950, a eletrônica foi utilizada para rastrear animais selvagens e que na década de 1970 foi realizada uma das primeiras tentativas de observação do comportamento do gado com a eletrônica. O primeiro trabalho para rastrear animais utilizando GNSS (*Global Navigation Satellite System* - Sistema Global de Navegação por Satélite) foi registrado em 1991, utilizando GPS. O monitoramento do gado, utilizando GPS, foi registrado pela primeira vez em 1997 ([ANDERSON; CIBILS, 2013](#)).

Diversos trabalhos desenvolvidos entre 1997 e 2002, utilizando dispositivos de GNSS para monitorar o comportamento do gado, são apresentados por [Anderson e Cibils \(2013\)](#). A tecnologia GNSS combinada com outros equipamentos eletrônicos é utilizada com frequência para monitorar os animais e seus comportamentos.

[Watanabe et al. \(2005\)](#) cita trabalhos que fazem uso de acelerômetros para medir os sinais do movimento e da postura do corpo em diferentes animais. O trabalho, que monitorou um gato doméstico, indica que os dados de aceleração podem ser úteis para

indicar as atividades e o comportamento de animais terrestres. No caso de herbívoros, os dados poderiam ser utilizados para indicar a quantidade de grama ingerida pelo animal, por exemplo. Capturar os movimentos de animais livres requer pequenos sensores e o armazenamento de grandes conjuntos de dados. Comportamentos básicos normalmente envolvem movimentos de todo o corpo: postura do corpo, sentidos de movimento característicos e as frequências de movimentos distintos (SCHEIBE; GROMANN, 2006). A partir dos dados de aceleração, deve ser possível identificar os diferentes comportamentos por meio da análise do padrão de aceleração dos movimentos do animal.

As soluções comerciais disponíveis no mercado para rastrear animais, em especial da tecnologia GNSS, são caras. As opções de baixo custo normalmente são soluções criadas a partir de dispositivos comerciais, não projetadas para o monitoramento animal. Existem vários trabalhos que adaptaram essas soluções com sucesso, incluindo o trabalho de Jesus (2014) que utiliza um módulo GPS conectado à plataforma Arduino.

A utilização de dispositivos não comerciais pode apresentar outras vantagens, como a possibilidade de inclusão de outros recursos que possam avaliar novas variáveis relacionadas ao comportamento animal e o menor tempo de manutenção, em caso de falhas do equipamento, considerando que não há dependência do fabricante para a solução (ANDERSON; CIBILS, 2013). Porém, há a necessidade de alguém localmente capacitado para que se ofereça suporte.

Em geral, as soluções para identificação do comportamento animal, utilizando recursos computacionais, são limitadas pelo consumo de energia e a capacidade da bateria. Para aplicações no Agronegócio, em especial de monitoramento do gado (HUIRCÁN *et al.*, 2010), é fundamental que o sistema possua recursos de baixo custo, em função do elevado número de elementos e uma gestão de energia eficiente, para manter o sistema o maior tempo disponível (considerando a dificuldade de acesso aos nós).

A tecnologia de RSSF é uma candidata natural para este tipo de aplicação (NADIMI *et al.*, 2008), pelo fato de incorporar em um único sistema a capacidade de sensoriamento, processamento e comunicação por uma rede sem fio, permitindo que os dados sejam coletados e encaminhados regularmente para um sistema central de controle e gestão. As RSSFs foram utilizadas para a identificação do comportamento e da localização dos animais em Huircán *et al.* (2010); Nadimi *et al.* (2008); Nadimi, Søggaard e Bak (2008); Nadimi *et al.* (2011); Nadimi *et al.* (2012); Silva *et al.* (2005).

2.3.4 Instrumentos para Coleta de Dados

Um dos métodos mais comum para equipar os animais, é a utilização de colares. O colar não deve ficar apertado nem solto no pescoço do animal, pois apertado pode abrasonar a pele e solto pode enroscar em algum objeto do ambiente (galho, cerca,

etc) (ANDERSON; CIBILS, 2013). Além disso, alguns componentes precisam ficar em posição fixa, como antenas ou sensores, pois se o colar se movimentar no pescoço do animal a qualidade dos dados coletados será afetada.

Algumas alternativas citadas por Anderson e Cibils (2013), a fim de evitar que o colar deslize sobre o pescoço do animal, são: deixar os componentes mais pesados (por exemplo, as baterias) na parte inferior do colar, a fim de agirem como contrapeso; ou um colar composto por material elástico, que absorva o suor e se estenda pelo pescoço do animal sem abrasaná-lo.

Um dos desafios, ao colocar equipamentos em animais livres na natureza, é garantir que os comportamentos do animal não sejam adversamente afetados. O *design* e o tamanho do dispositivo devem ser projetados de maneira que possua o menor peso e tamanho possíveis. Watanabe *et al.* (2005) indicam que o equipamento não deve ter mais que 5% do peso do animal enquanto Brooks e Harris (2008 apud ANDERSON; CIBILS, 2013) recomendam que o peso seja menor que 1% da massa corporal do animal. Após equipar o animal, deve-se aguardar um período para que o mesmo se adapte ao equipamento. Não há um tempo ideal de espera indicado, pois é variável conforme a espécie avaliada. Por exemplo, em ovinos um período de 16 horas, entre a instrumentação e o início da coleta dos dados, parece ser ideal, segundo Hulbert *et al.* (1998 apud ANDERSON; CIBILS, 2013).

Em geral, Scheibe e Gromann (2006) apontam que o sistema deve ser pequeno, leve, resistente a choques e simples para fixar ao corpo do animal.

2.3.5 Influência do Ambiente no Comportamento Animal

Sabe-se que fatores ambientais relacionam-se diretamente com o comportamento animal. Oliveira (2013a) elenca a temperatura, a radiação solar e a umidade relativa como os principais elementos do ambiente relacionados ao comportamento bovino. Segundo Schütz *et al.* (2009 apud OLIVEIRA, 2013a), o gado pode identificar pequenas alterações climáticas e buscar situações para contorná-las, como procurar áreas com sombra para reduzir a temperatura corporal. No trabalho de Kilgour *et al.* (2012), identificou-se que 10,9% do tempo os animais estavam em áreas de sombra.

2.4 Reconhecimento de Padrões

No cenário mundial, globalizado, dinâmico e concorrido, a tomada de decisão é cada vez mais difícil e arriscada. Com o volume de dados gerados nos diferentes setores, a interpretação deles fica cada vez mais complexa, e em alguns casos humanamente impossíveis de serem realizadas. O reconhecimento de padrões é uma solução para o problema. Para Theodoridis e Koutroumbas (2006), o reconhecimento de padrões tem como objetivo

a classificação de objetos (padrões) em um número de categorias ou classes, através de algoritmos computacionais. Reconhecimento de padrões, em geral, baseia-se na extração de características capazes de discriminar objetos em diferentes classes. O processo de reconhecimento de padrão é dividido basicamente em duas fases: treinamento (aprendizagem) e reconhecimento (classificação).

O Reconhecimento de Padrões – subtópico da Aprendizagem de Máquina, que por sua vez, é um subcampo da Inteligência Artificial – é definido por Duda, Hart e Stork (2000) como o ato de tomar uma decisão a partir de um conjunto de dados, baseado em um padrão. Os padrões são modelos que servem de referência e que representam a relação entre os objetos avaliados.

A identificação de padrões de maneira automatizada é uma atividade complexa mas imensamente útil. A partir da categorização dos dados procura-se encontrar a hipótese mais provável dentro de um conjunto de hipóteses. Para Duda, Hart e Stork (2000) é natural que se busque projetar e construir máquinas que reconheçam padrões baseado nos sistemas de reconhecimento de padrões do mundo natural. Na natureza os padrões se manifestam de diversas maneiras, sons, formas, imagens, cheiros, e a todo instante os seres humanos e os animais, percebem e interagem com estes padrões. Por exemplo, baseado em padrões, o homem diferencia o som de uma música e de um motor e animais diferenciam a presa do predador.

As possibilidades de aplicações de reconhecimento de padrões são várias, entre elas: a identificação de impressões digitais, reconhecimento óptico de caracteres, análise de sequências de DNA, reconhecimento de alvos militares. Para algumas aplicações, os esforços para encontrar um padrão podem ser influenciados pelo conhecimento de como estes são resolvidos na natureza, tanto nos algoritmos quanto no modelo de *hardware* proposto. Sendo assim, o conhecimento prévio sobre o problema é importante para o desenvolvimento das aplicações, sendo fundamental a presença de especialistas do domínio no desenvolvimento da aplicação.

2.4.1 Mineração de Dados

A Mineração de Dados é uma área de pesquisa da Inteligência Artificial que busca extrair conhecimentos a partir de bases de dados (*Knowledge Discovery in Database - KDD*). O processo geral de conversão de dados brutos em informações úteis, consiste em uma série de passos de transformação, pré-processamento dos dados e pós-processamento dos resultados.

Segundo Tan, Steinbach e Kumar (2009), o pré-processamento consiste em transformar os dados brutos em um formato apropriado para análises subsequentes, incluindo a fusão de dados de diferentes fontes, remoção de ruídos, de observações duplicadas e

seleção de registros e características relevantes à tarefa de mineração de dados. O pós-processamento é a etapa de análise dos resultados válidos e úteis para incorporar o sistema de apoio a decisões. A Figura 3 ilustra o processo de descoberta de conhecimento em base de dados.

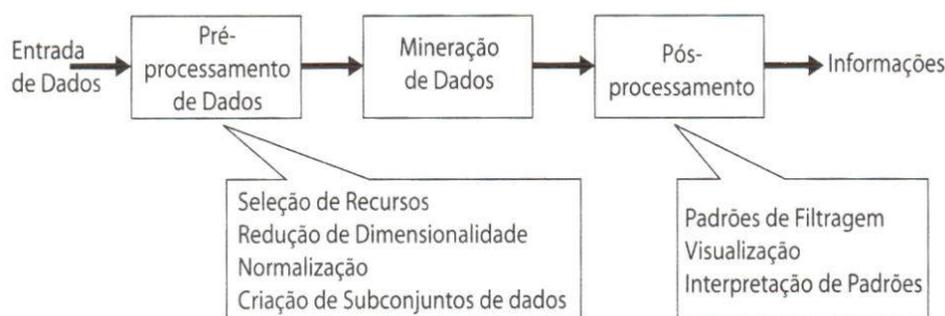


Figura 3 – Processo de descoberta de conhecimento (TAN; STEINBACH; KUMAR, 2009).

As tarefas de Mineração de Dados são divididas em duas categorias (TAN; STEINBACH; KUMAR, 2009):

- Tarefas de Previsão: utilizadas para prever o valor de um atributo a partir dos valores de outros atributos.
- Tarefas Descritivas: utilizadas para derivar padrões (correlações, tendências, grupos, trajetórias e anomalias) que resumem os relacionamentos subjacentes nos dados.

A aprendizagem do modelo pode ser feita de maneira supervisionada ou não supervisionada. Na aprendizagem supervisionada, um rótulo da categoria é fornecido para cada padrão em um conjunto de treinamento e o algoritmo busca aprender uma solução para rotular instâncias desconhecidas. Já na aprendizagem não supervisionada, o sistema forma *clusters* ou agrupamentos dos padrões de entrada, e o modelo deve descobrir sozinho as relações, padrões ou categorias nos dados de entrada (DUDA; HART; STORK, 2000).

2.4.2 Modelos de Classificação

Tan, Steinbach e Kumar (2009) definem a classificação como a tarefa de rotular cada registro em uma das categorias pré-definidas. Cada registro é caracterizado por uma tupla (x, y) , onde x é o conjunto de atributos e y o atributo especial, chamado de atributo alvo ou de categorização. Uma função alvo (modelo) mapeia cada conjunto de atributos x para um dos rótulos de classes y pré-determinados.

Para construir um modelo é necessário que um conjunto de dados, com o atributo alvo rotulado e denominado **conjunto de treinamento**, seja processado com um

algoritmo de classificação. O modelo então é validado com um **conjunto de teste**, com registros com rótulos de classes desconhecidos. A Figura 4 apresenta a abordagem geral para a construção de um modelo de classificação (TAN; STEINBACH; KUMAR, 2009).

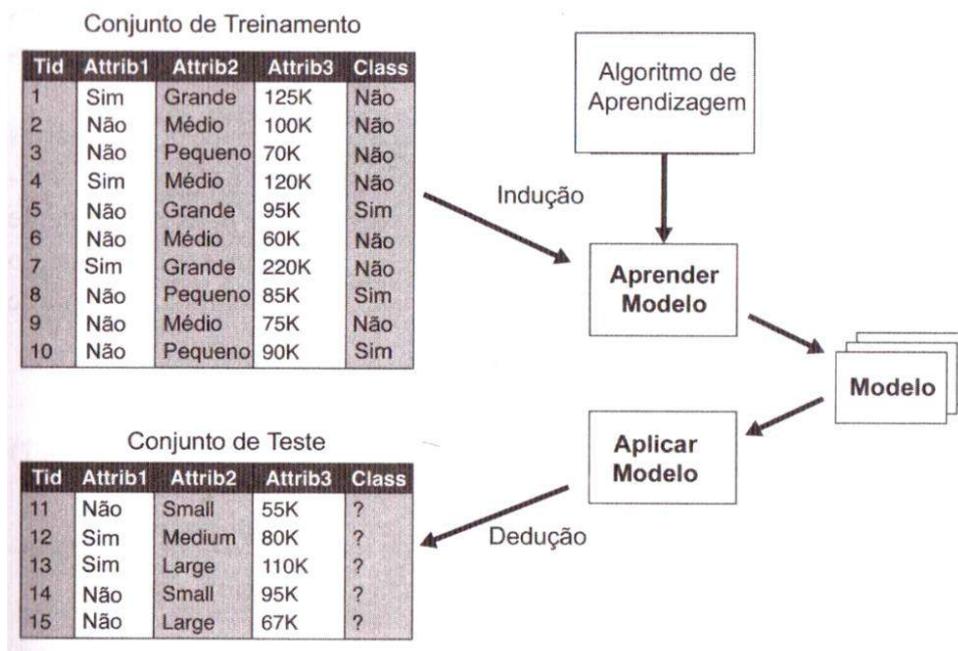


Figura 4 – Abordagem geral para a construção de um modelo de classificação (TAN; STEINBACH; KUMAR, 2009).

2.4.3 Técnicas para a Construção de Modelos de Classificação

Diferentes algoritmos de aprendizado de máquina estão disponíveis para a tarefa de classificação, distribuídos nas categorias Indução de Árvores de Decisão, Baseados em Regras, Redes Neurais, Máquinas de Vetor de Suporte, Classificadores Bayesianos, Métodos de Grupos, entre outros. A seguir apresenta-se sucintamente algumas destas categorias.

- **Indução de Árvores de Decisão:** uma árvore de decisão é uma estrutura hierárquica dos dados, baseada em estágios de decisão (nós) e separação de classes. Ela é formada por um nó inicial chamado de raiz, de nós internos e nós folhas. O nó raiz e os nós internos são considerados não terminais, possuem duas ou mais ramificações e condições de testes de um ou mais atributos, enquanto o nó folha possui um rótulo e nenhuma ramificação (TAN; STEINBACH; KUMAR, 2009).

As diferentes ligações a partir de um nó, correspondem as possíveis respostas às condições de teste. As ligações devem ser mutuamente distintas e somente uma será seguida. A ligação para o nó subsequente é seguida, continuando até encontrar um

nó folha, que apresenta o rótulo para categoriar o registro (DUDA; HART; STORK, 2000).

Duda, Hart e Stork (2000) destacam o uso de árvores de decisão, comparado a outras técnicas, pela sua interpretabilidade. No exemplo da Figura 5 é possível identificar que um registro com os atributos (Cor = Amarelo e Formato = Fino) é rotulado como Banana; e que o rótulo Maçã pode ser atribuído aos elementos com os atributos (Cor = Verde e Tamanho = Médio) ou (Cor = Vermelho e Tamanho = Médio).

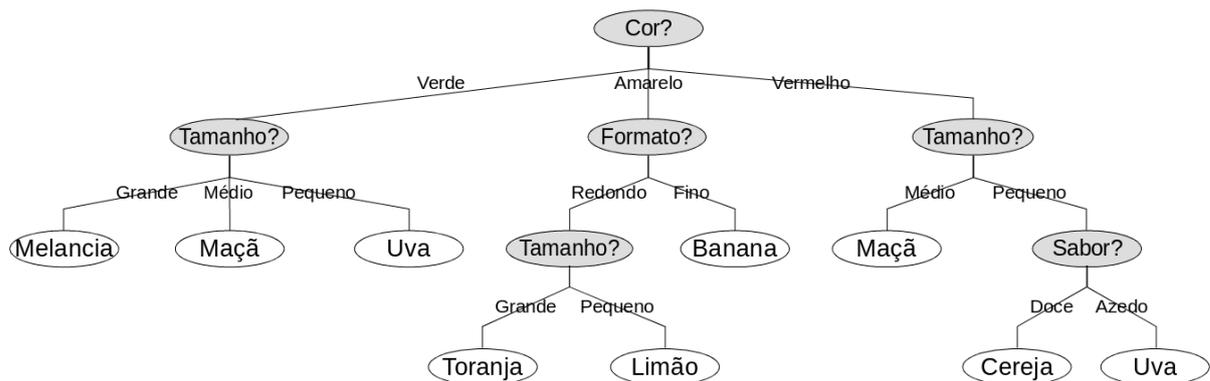


Figura 5 – Exemplo de árvore de decisão: classificação de frutas baseado no sabor, cor, formato e tamanho (DUDA; HART; STORK, 2000).

- **Classificadores Baseados em Regras:** é uma técnica que utiliza um conjunto de regras "se ... então" para classificar os registros. Uma vez que a *condição* é satisfeita, o caso então é identificado com a *classe*.

A expressividade de um conjunto de regras é quase equivalente ao de uma árvore de decisão, já que cada nó da árvore gera um regra. Segundo Tan, Steinbach e Kumar (2009), os classificadores baseados em regras permitem que múltiplas regras sejam disparadas para um determinado registro, permitindo a construção de um modelo com limite de decisão mais complexo. Uma abordagem utilizada por esta técnica é a ordenação das classes, que se mostra bem apropriada para trabalhar com conjuntos de dados com distribuições de classes desequilibradas.

- **Rede Neural Artificial:** inspirada na estrutura do cérebro humano, é composta de um conjunto de interconexões de nós (neurônios) e ligações direcionadas (axônios). Os nós de entrada possuem os sinais de entradas, que são multiplicados por pesos, que indicam a sua influência no nó de saída. Cada modelo de rede neural possui uma regra de treinamento, onde os pesos de suas conexões são ajustados de acordo com os padrões apresentados, baseado nos exemplos de entrada. Entre os nós de

entrada e saída, há os nós intermediários (ocultos), onde é feito o processamento através das conexões. A Figura 6 mostra um exemplo de rede neural artificial.

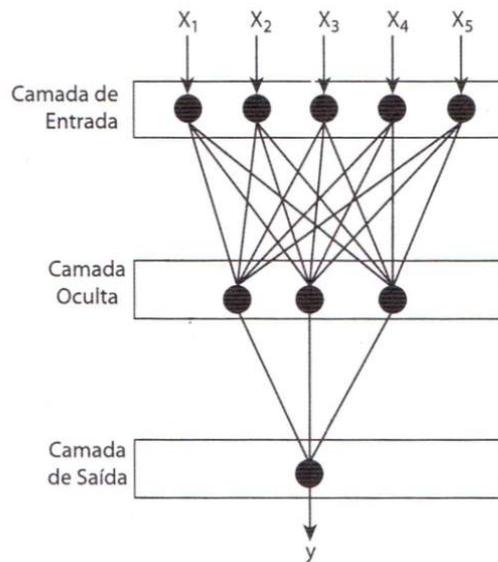


Figura 6 – Exemplo de uma rede neural artificial multicamadas (TAN; STEINBACH; KUMAR, 2009).

- Métodos de Grupos:** esta técnica agrega múltiplos classificadores para melhorar a precisão das classificações. Um método constrói um conjunto de classificadores a partir dos dados de treinamento e executa a classificação recebendo um voto sobre as previsões feitas por cada um dos classificadores. A Figura 7 apresenta a ideia de se criar múltiplos classificadores e agregar suas previsões ao classificar os registros. Tan, Steinbach e Kumar (2009) apresentam os conceitos e os detalhes dos algoritmos *Bagging*, *Boosting* e *Random Forest*, exemplos de algoritmos que implementam esta técnica.

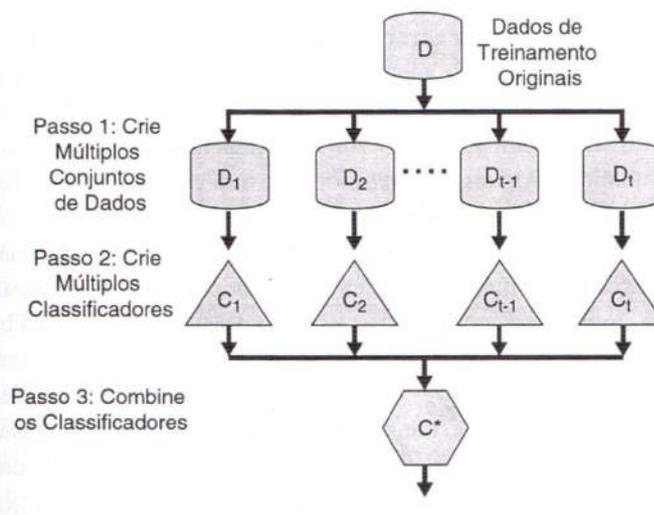


Figura 7 – Visão lógica do método de aprendizagem de grupo (TAN; STEINBACH; KUMAR, 2009).

2.4.4 Overfitting de Modelo

Os erros cometidos por um classificador são divididos em dois tipos: erros de treinamento e erros de generalização (TAN; STEINBACH; KUMAR, 2009). Um bom modelo deve ter baixa quantidade de erros de treinamento assim como de generalização, ou seja, não deve se adaptar apenas aos dados de treinamento, mas também classificar com precisão os registros nunca vistos antes.

Quando o modelo se ajusta excessivamente aos dados do conjunto de treinamento (regras muito especializadas) e exemplos desconhecidos são classificados de maneira errada, temos o chamado *overfitting* de modelo. Quando o modelo tem um desempenho fraco no conjunto de treinamento, em geral porque ele não conhece a verdadeira estrutura dos dados, temos o chamado *underfitting* de modelo.

Estes problemas acontecem pelo desbalanceamento das classes (existência de classe majoritária) ou pela divisão do conjunto de dados para treinamento e teste, em que a seleção dos dados para treinamento não contemplam as características de todo o conjunto de dados.

2.4.5 Seleção de Atributos (Dimensionalidade)

A dimensionalidade está relacionada ao volume de dados utilizado por um método de aprendizagem. A redução da dimensionalidade faz com que muitos algoritmos de mineração de dados funcionem melhor, pois ela pode eliminar características irrelevantes, além de que a análise de dados se torna mais difícil quando a dimensionalidade dos dados aumenta. A redução do número de atributos ou a seleção dos atributos mais relevantes são técnicas utilizadas para reduzir a dimensionalidade dos dados. O processo de seleção de atributos simplifica o modelo de predição, eliminando atributos redundantes e irrelevantes, com o objetivo de melhorar a qualidade das predições realizadas (TAN; STEINBACH; KUMAR, 2009).

A abordagem ideal para a seleção de atributos é experimentar todos os subconjuntos possíveis como entrada para o algoritmo de interesse, porém, isto não é possível na maioria das situações, em função do custo computacional para combinar todas as possibilidades.

A maioria dos algoritmos de aprendizagem de máquina já faz a seleção dos atributos durante a sua execução, decidindo quais atributos usar e quais ignorar. Porém, nem sempre essa seleção ocorre efetivamente e há utilização de atributos que influenciam negativamente. Assim, estratégias alternativas são utilizadas para melhorar a seleção. Além da abordagem de seleção interna, realizada pelos próprios algoritmos de aprendizagem de máquina, há outras duas apontadas por Tan, Steinbach e Kumar (2009): Abordagem por Filtros e Abordagem *Wrappers*.

Uma arquitetura comum para o processo de seleção é proposta para ambas as abordagens (Figura 8). O processo consiste em quatro etapas: uma medição para avaliar um subconjunto; uma estratégia de pesquisa que controla a geração de um novo subconjunto; um critério de parada e um procedimento de validação.

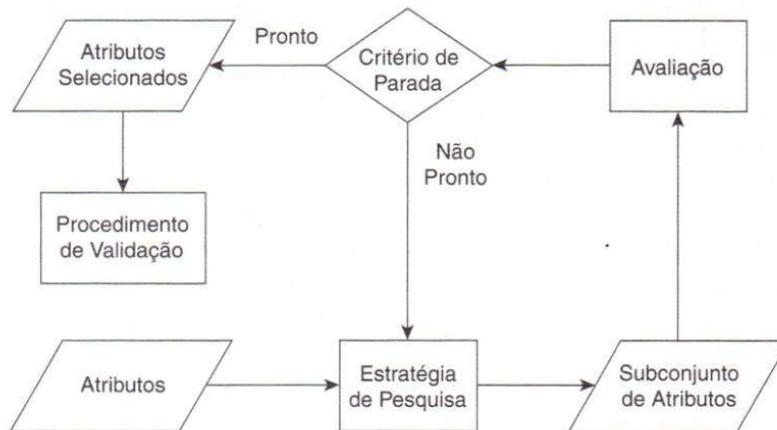


Figura 8 – Diagrama de fluxo de um processo de seleção de subconjuntos de características (TAN; STEINBACH; KUMAR, 2009).

A diferença da abordagem por Filtros e *Wrappers* é a forma de avaliar o subconjunto. A abordagem *Wrappers* utiliza o algoritmo de aprendizagem de máquina como uma caixa preta, para encontrar o melhor subconjunto de atributos, gerando um subconjunto baseado em um algoritmo específico. Na abordagem por Filtros os atributos são selecionados antes da execução do algoritmo, ou seja, a seleção é independente do algoritmo utilizado.

2.4.6 Metodologia para Avaliação de Modelos

No momento da construção do modelo o algoritmo só possui acesso ao conjunto de treinamento, desconhecendo o conjunto de teste. Assim, é necessário avaliar o erro de generalização. Três métodos, baseado em Tan, Steinbach e Kumar (2009), são apresentados a seguir:

- **Re-substituição:** Esta abordagem considera que o conjunto de treinamento é uma boa representação dos dados gerais e pode ser usado para fornecer uma avaliação otimista para o erro de generalização. Entretanto, o erro de treinamento geralmente é uma estimativa fraca do erro de generalização.
- **Holdout:** Os dados são divididos em dois conjuntos, chamados de conjunto de treinamento e conjunto de teste. O modelo então é induzido a partir do conjunto de treinamento e seu desempenho avaliado pelo conjunto de teste. Este método

possui como limitações o fato do conjunto ser fracionado, reduzindo o número de registros utilizados no treinamento. Quanto menor o conjunto de treinamento menor a variância, em contra partida quanto maior o conjunto de treinamento maior a confiança no conjunto de teste.

- **Cross-Validation:** O conjunto é dividido em n partes (*fold*s) de tamanhos aproximados, nos quais $n-1$ blocos são utilizados como conjunto de treinamento e a parte restante como conjunto de teste. O processo repete-se n vezes, até que todos os *fold*s sejam utilizados como conjunto de teste. O erro total é a média dos erros de todas as execuções.

2.4.7 Métricas para Avaliação de Modelos

A matriz de confusão apresenta detalhes do resultado de validação de um modelo, detalhando os erros e acertos do classificador para cada uma das classes. Ela é uma matriz quadrada $N \times N$, onde N é o número de classes investigadas. As linhas representam as classes investigadas e as colunas as associações definidas pelo classificador.

Os elementos da matriz diagonal representam o número de exemplos corretamente classificados, identificados como Verdadeiro Positivo (VP) e Verdadeiro Negativo (VN). Os exemplos classificados incorretamente são identificados como Falsos Positivos (FP) e Falsos Negativos (FN), sendo Positivo e Negativo exemplos de classes. A Tabela 1 mostra um exemplo de matriz com duas classes, em que a Classe A possui 57 instâncias rotuladas corretamente (VP) e 3 incorretamente (FN - classificadas como Classe B), e a Classe B possui 15 instâncias rotuladas incorretamente (FP - classificadas como Classe A) e 23 classificadas corretamente (VN).

Tabela 1 – Exemplo de matriz de confusão.

		Classe Prevista	
		Classe A	Classe B
Classe Real	Classe A	57 (VP)	3 (FN)
	Classe B	15 (FP)	23 (VN)

Para comparar os resultados de dois modelos avaliados, é mais conveniente resumir as informações da tabela de confusão em um único número. Entre as métricas de desempenho para representar o quão bom é um modelo, temos:

- **Taxa de Acerto:** razão entre a quantidade de instâncias corretamente classificadas e o total de instâncias.

$$\frac{\text{Número de Previsões Corretas}}{\text{Número Total de Previsões}}$$

- **Taxa de Erro:** razão entre a quantidade de instâncias classificadas incorretamente e o total de instâncias; ou simplesmente a diferença entre 1 e a taxa de acerto.

$$\frac{\text{Número de Previsões Erradas}}{\text{Número Total de Previsões}}$$

ou

$$\text{Taxa de Erro} = 1 - (\text{Taxa de Acerto})$$

As aplicações reais, em geral, apresentam conjuntos de dados com distribuições de classes desequilibradas. A medida de precisão, que é extensivamente utilizada para comparar o desempenho dos classificadores, pode não ser bem apropriada para avaliar modelos derivados de conjuntos de dados desequilibrados (TAN; STEINBACH; KUMAR, 2009). Por exemplo, se 1% das transações com cartões de crédito forem fraudulentas e o modelo prever que toda as transações são legítimas, há uma taxa de acerto de 99%.

Métricas alternativas são utilizadas para avaliar as classes de um modelo. Algumas delas são:

- **Taxa de Verdadeiros Positivos (TVP) ou Sensibilidade:** avalia a capacidade do modelo classificar uma instância como Positivo, quando realmente ela é Positivo. É a fração dos exemplos positivos previstos corretamente pelo modelo.

$$\frac{VP}{(VP+FN)}$$

- **Taxa de Verdadeiros Negativos (TVN) ou Especificidade:** avalia a capacidade do modelo predizer um registro como Negativo, quando realmente ele é Negativo. É a fração dos exemplos negativos previstos corretamente.

$$\frac{VN}{(VN+FP)}$$

A sensibilidade e a especificidade variam em direções opostas, isto é, se o modelo é muito sensível a positivos, tende a gerar muitos falso-positivos, e vice-versa. Assim, um método de decisão perfeito (100% de sensibilidade e 100% especificidade) raramente é alcançado, e um balanço entre ambos deve ser atingido.

De acordo com Tan, Steinbach e Kumar (2009), a precisão e a cobertura são duas métricas amplamente usadas em aplicações onde a detecção bem sucedida de uma das classes é considerada mais significativa do que a de outras classes.

- **Precisão (*Precision*):** é o valor do número de casos positivos por total de casos cobertos, muito influenciada pela especificidade.

$$\frac{VP}{(VP+FP)}$$

- **Cobertura (*Recall*):** é influenciado pela sensibilidade e pouco pela especificidade. É calculada por número de casos cobertos pelo número total de casos aplicáveis.

$$\frac{VP}{(VP+FN)}$$

- **Medida-F (*F-measure*):** é uma média harmônica entre a Precisão e a Cobertura. Um valor alto da Medida-F assegura que tanto a Precisão quanto a Cobertura são razoavelmente altas. Também conhecido como *F-Score* ou F1.

$$\frac{2((Precisão)(Cobertura))}{Precisão+Cobertura}$$

- **Índice Kappa:** compara o valor encontrado nas observações com aquele que se pode esperar do acaso. A Tabela 2, proposta por Landis e Koch (1977), apresenta as faixas de valores do índice *Kappa*, onde o valor 1 implica a correlação perfeita. Quanto menor o valor de *Kappa*, menor a confiança da observação.

Tabela 2 – Interpretação dos valores do índice Kappa (LANDIS; KOCH, 1977).

Valores de Kappa	Nível de Concordância
< 0.00	Nenhuma
0.00-0.20	Pouca
0.21-0.40	Mediana
0.41-0.60	Moderada
0.61-0.80	Substancial
0.81-1.00	Quase perfeita

2.4.8 Ferramenta WEKA

O pacote WEKA (*Waikato Environment for Knowledge Analysis*) é uma coleção de algoritmos de aprendizado de máquina, de código aberto, que pode ser aplicado para tarefas de mineração de dados. Ele contém ferramentas para pré-processamento de dados, classificação, regressão, *clustering*, regras de associação e visualização (WEKA, 2015).

Ele é desenvolvido na linguagem Java, contém uma GUI (*Graphical User Interface*) para interagir com arquivos de dados e produzir resultados visuais. A Figura 9 apresenta a tela inicial do WEKA, com as opções para utilização da interface gráfica (*Explorer*) e linha de comando (*Simple CLI*). Também é possível usar as suas funcionalidades a partir da sua API (*Application Programming Interface*) em um código Java.

A interface gráfica contém as funcionalidades de pré-processamento, análise (classificação, associação e clusterização) e visualização dos resultados através de janelas. A Figura 10 apresenta a tela de pré-processamento (*Preprocess*) na interface gráfica, onde o usuário pode selecionar o arquivo com o conjunto de dados de entrada e visualizar estatísticas do conjunto selecionado, como o número de atributos e o número de instâncias.



Figura 9 – Tela inicial do WEKA.

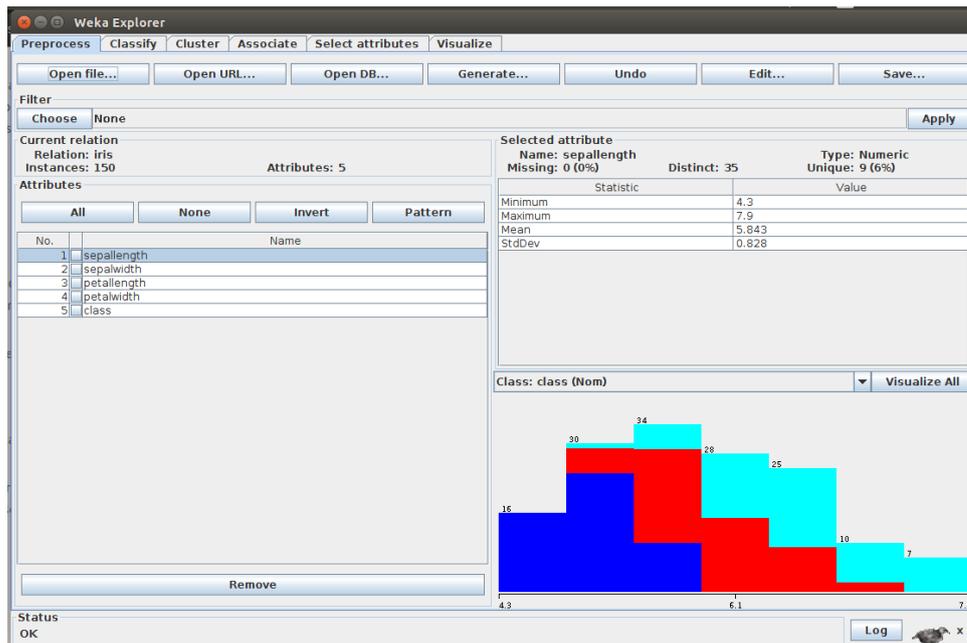


Figura 10 – Tela de pré-processamento do WEKA.

A Figura 11 mostra a aba classificadores (*Classify*), com o algoritmo *Random Forest* selecionado. Na opção *Test options* o usuário seleciona o método de validação desejado. Há quatro modos:

- **Use training set:** testa o classificador com o próprio conjunto de treinamento utilizado na construção do modelo.
- **Supplied test set:** o usuário pode selecionar um outro conjunto de dados (outro arquivo) para validar o classificador.
- **Cross-validation:** para validação com o método *cross-validation* com a opção de seleção da quantidade de partes (*folds*) que o conjunto será dividido.
- **Percentage split:** divide o conjunto de dados em duas partes, uma para treinamento do modelo e outra para validação dele. No campo percentual indica-se a parte dos registros utilizados na construção do modelo.

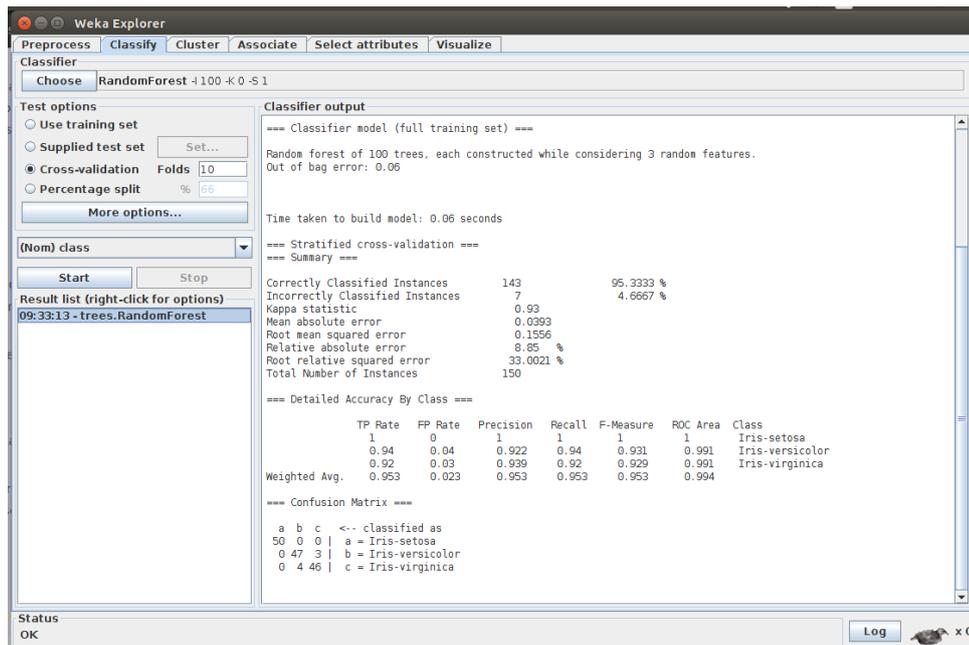


Figura 11 – Tela com as opções do método de classificação do WEKA.

O WEKA é reconhecido como um sistema de referência em mineração de dados e aprendizado de máquina, com ampla aceitação nos meios acadêmicos e de negócios, tornando-se uma ferramenta amplamente utilizada em mineração de dados (HALL *et al.*, 2009).

2.5 Considerações Finais

Este capítulo apresentou um breve panorama dos campos de estudo abordados pelo trabalho, buscando introduzir os principais conceitos explorados pela pesquisa, sem pretensões de esgotá-los.

Apesar da abordagem sobre as redes de sensores, este trabalho não implementou a comunicação entre os nós da rede. O envio dos dados para a estação base foi realizada manualmente, através de um cartão de memória, que armazenou os dados coletados em cada nó. Isso não descaracteriza o uso delas, já que outras características das RSSFs são relevantes para as abordagens com sensores.

As principais características sobre comportamento animal foram apresentadas, apontando as oportunidades e dificuldades de exploração da área, assim como os sensores utilizados, apresentando suas características perante a pesquisa.

Por fim, no campo de reconhecimento de padrões, as técnicas e algoritmos de classificação foram detalhados, apresentando suas características para a efetivação do objetivo principal do trabalho: a identificação dos comportamentos bovinos.

3 Trabalhos Relacionados

As pesquisas visando automatizar o monitoramento do comportamento animal têm se tornado cada vez mais atraentes, tanto para fins científicos quanto para fins produtivos. Neste capítulo, são apresentadas as características dos trabalhos que serviram de base e fonte de informações para o desenvolvimento desta pesquisa, mostrando as abordagens e técnicas utilizadas por estes para automatizar a identificação do comportamento animal. Na Seção 3.1, são apresentados os equipamentos e sensores utilizados nas pesquisas. A Seção 3.2 e Seção 3.3 apresentam os detalhes do processo de coleta de dados e dos comportamentos observados. Na Seção 3.4, as técnicas utilizadas para o tratamento e processamento dos dados. Por fim, na Seção 3.5, é mostrado um resumo dos resultados apresentados pelos trabalhos.

3.1 Equipamentos e Sensores Utilizados

Em geral, os equipamentos são colocados nos animais em colares, que ficam presos no pescoço do animal. Destaque para os trabalhos de [Moreau *et al.* \(2009\)](#), [Scheibe e Gromann \(2006\)](#), [Martiskainen *et al.* \(2009\)](#) e [Robert *et al.* \(2009\)](#).

[Moreau *et al.* \(2009\)](#) coletaram dados com equipamentos colocados simultaneamente em três lugares do corpo de cabras: um no pescoço, um na intersecção do dorso/pescoço e outro nas costas do animal. Foram comparados os resultados de um classificador utilizando os dados de um sensor diferente daquele utilizado na construção do modelo de classificação. Os resultados indicaram influência significativa da posição do colar utilizado, afirmando que se deve utilizar os dados do mesmo equipamento para classificar o comportamento. [Scheibe e Gromann \(2006\)](#) também colocou o equipamento em partes diferentes do corpo do animal (um colar preso no pescoço e outro na perna), com o objetivo de aferir os comportamentos de vacas e cavalos, e comparar os resultados de cada um dos sensores. Já [Robert *et al.* \(2009\)](#) utilizou uma abordagem diferente, prendendo o equipamento apenas na perna do animal, já que os comportamentos de interesse da pesquisa (Em Pé, Deitado e Andando) estão relacionados ao posicionamento da perna.

[Martiskainen *et al.* \(2009\)](#) apontaram que pequenas variações no posicionamento do equipamento no animal influenciam os resultados da classificação e que a padronização da inclinação do acelerômetro pode melhorar o desempenho do classificador, reforçando [Watanabe *et al.* \(2005\)](#), que indicaram a necessidade de calibração dos sensores de aceleração, individualmente para cada animal avaliado.

Nos trabalhos avaliados, os principais sensores utilizados foram os sensores de

movimentação, com destaque para o acelerômetro tridimensional, que foi utilizado nos experimentos de Scheibe e Gromann (2006), Martiskainen *et al.* (2009), Robert *et al.* (2009) e Diosdado *et al.* (2015). Outro dispositivo comumente utilizado é o GPS. Oliveira (2013b) e Jesus (2014) usaram exclusivamente este dispositivo, enquanto Moreau *et al.* (2009), González *et al.* (2015) e Dutta *et al.* (2015) utilizaram o GPS e o acelerômetro. Guo *et al.* (2009) utilizaram, além do acelerômetro e GPS, o giroscópio.

Em geral, o acelerômetro e o GPS foram configurados para coleta de dados em uma frequência de 10Hz e 4Hz, respectivamente. O trabalho de Robert *et al.* (2009) foi o único avaliado que utilizou uma frequência diferente para o acelerômetro, 100Hz.

3.2 Coleta de Dados

Diosdado *et al.* (2015), Dutta *et al.* (2015), González *et al.* (2015), Guo *et al.* (2009), Martiskainen *et al.* (2009), Robert *et al.* (2009) e Scheibe e Gromann (2006) monitoraram os comportamentos de bovinos, enquanto Moreau *et al.* (2009) monitorou a atividade de cabras e Watanabe *et al.* (2005) de gatos domésticos. O número de animais equipados nos experimentos e o tempo de coleta de dados variou entre os trabalhos. Martiskainen *et al.* (2009) coletou dados de 30 animais por um período de 30 dias, enquanto Diosdado *et al.* (2015) monitorou 6 animais por um período de apenas 36 horas.

Moreau *et al.* (2009) realizaram duas coletas, em lugares geograficamente diferentes (um na Europa e outro na Ásia) e com animais de raças diferentes, com o objetivo de avaliar se as características do animal e do sistema de criação (tipo de terreno e de pastejo) têm efeito sobre a qualidade da classificação automatizada dos comportamentos. Além disso, o intervalo de coleta dos dados foi diferente no dois experimentos, a cada 1 segundo no primeiro e a cada 2 segundos no segundo. Robert *et al.* (2009) também dividiu as coletas em intervalos diferentes, dividindo os 15 bezerros monitorados em três grupos, configurados para coleta nos intervalos de 1, 5 e 10 segundos.

Duas técnicas, para armazenar e recuperar os dados coletados, foram encontradas nos trabalhos avaliados: armazenados em uma memória interna, para posterior recuperação, quando o animal é novamente capturado; ou transmitidos, por uma rede sem fio, para um receptor central. Krause *et al.* (2013 apud DIOSDADO *et al.*, 2015) definem estes dois métodos de recuperação dos dados, respectivamente, como *bio-loggers* e *bio-telemetry sensors*.

3.3 Comportamentos Observados

Os comportamentos mais comumente classificados nos trabalhos são Pastando/Procurando, Em Pé, Deitado e Andando. Alguns trabalhos abordam também o compor-

tamento Ruminando ((DUTTA *et al.*, 2015), (GONZÁLEZ *et al.*, 2015) e (MARTISKAINEN *et al.*, 2009)). Robert *et al.* (2009) abordou apenas os comportamentos Em Pé, Deitado e Andando, enquanto que Moreau *et al.* (2009), os comportamentos Andando, Parado e Pastando/Procurando.

Outra abordagem utilizada é categorizar aqueles comportamentos que não são de interesse em uma categoria geral. Por exemplo, se o trabalho não tinha interesse em avaliar o comportamento Bebendo Água, quando este acontecia era classificado como "Outros". Dutta *et al.* (2015) e González *et al.* (2015) utilizaram esta abordagem.

Martiskainen *et al.* (2009) destacam a necessidade de trabalhos que busquem identificar um número maior de comportamentos dos animais, já que alguns dos sistemas propostos identificaram apenas um ou dois comportamentos satisfatoriamente, reforçado por Nadimi *et al.* (2012), que indicam que modelos para classificar mais comportamentos não alcançaram resultados satisfatórios.

Para Nadimi *et al.* (2012), a observação do comportamento animal pode trazer benefícios, a partir da análise de situações que afetam o bem estar animal: melhorar o bem-estar dos animais afeta positivamente a patologia animal e a resistência a doenças; enquanto que níveis inadequados de bem-estar animal pode afetar significativamente o crescimento, a reprodução e a taxa de sobrevivência dos animais, comprometendo a qualidade e a segurança dos alimentos produzidos. Um sistema de monitoramento pode ajudar o produtor a melhorar a gestão do negócio, aumentando o bem-estar animal e a produção.

Neste sentido, Scheibe e Gromann (2006) buscaram identificar comportamentos comuns (Andando, Correndo e Pastando), mas avaliando animais doentes e animais saudáveis, com o objetivo de verificar se há variações de movimentos entre eles. O resultado apontado foi que os animais doentes tiveram comportamentos identificados com características diferentes dos saudáveis, ou seja, que a partir dos dados de movimentação é possível diferenciar os comportamentos de animais doentes e saudáveis.

Nadimi *et al.* (2008) apontam que a identificação comportamental e da localização animal possibilitam que se avalie em que condições determinados comportamentos ocorrem, evoluindo no sentido de identificar individualmente animais com potenciais problemas de saúde ou otimizar o processo de pastejo destes. Porém, o foco dos trabalhos tem sido apenas indicar a localização do animal na área de pasto ou identificar os seus comportamentos. Moreau *et al.* (2009) apontam a possibilidade de combinar os dados do GPS com o acelerômetro para aferir os comportamentos e identificar as características dos comportamentos com o ambiente, porém utilizou os dados do GPS apenas para plotar em um mapa 3D a localização em que os comportamentos ocorreram.

Guo *et al.* (2009) utilizou os dados do acelerômetro e giroscópio para aferir os comportamentos (o trabalho não apresenta detalhes do uso destes sensores e nem de seus

resultados) e os dados do GPS para mostrar a relação das transições de localização com a mudança dos comportamentos, indicando as regiões em que os comportamentos ocorrem dentro da área de pasto. O trabalho de [Oliveira \(2013b\)](#) também avaliou a relação do animal com o ambiente, validando o uso de trajetórias semânticas para identificar as características do comportamento bovino. Utilizando um conjunto de dados georeferenciados e os dados do GPS, o trabalho relacionou o deslocamento do animal com as áreas de interesse (cocho, bebedouro, árvores) e apresentou informações sobre o deslocamento, como a distância percorrida e a duração e velocidade média das trajetórias.

3.4 Tratamento e Processamento dos Dados

Os conjuntos de dados utilizados para validar os modelos criados, em geral, são rotulados manualmente a partir de anotações realizadas em campo ou por análise de vídeos (gravados durante o experimento). Os dados registrados em manuscritos precisam ser transcritos para um arquivo digital, para então serem processados pelos algoritmos que fazem o reconhecimento dos comportamentos.

O trabalho [Dutta et al. \(2015\)](#) utilizou um *software* chamado *WhatISee*, desenvolvido com o objetivo de permitir anotações de comportamentos observados, registrando a data, horário e um rótulo, que identifica o comportamento observado. O *WhatISee* ainda possui a opção de notificação ao usuário, que é ativado após um intervalo de tempo pré-determinado. Trata-se de um aplicativo com função genérica, não desenvolvido especificamente para observação do comportamento animal, disponível para o sistema operacional *iOS*. [Jesus \(2014\)](#) desenvolveu o aplicativo *Ápis*, voltado especificamente para anotação de observações do comportamento animal, registrando a data, horário e o comportamento selecionado. A proposta do *software* é aperfeiçoar o trabalho de observação, importação e manipulação dos dados das anotações realizadas em campo. A versão está disponível para qualquer dispositivo móvel com o sistema operacional *Android*.

A utilização dos *softwares*, além de reduzir o tempo de rotulação dos dados, também elimina a possibilidade de erros na transcrição das anotações manuais para o arquivo digital.

Independente da técnica utilizada para rotular os registros, os conjuntos de dados foram organizados utilizando diferentes estratégias. Uma delas é a utilizada por [González et al. \(2015\)](#), onde dois conjuntos de dados foram gerados, sendo um para construir o modelo e outro para validá-lo. Abordagem também utilizada em [Moreau et al. \(2009\)](#), que utilizou o conjunto de dados de um animal para validar o modelo de classificação construído com os dados de outro animal.

Diferentes técnicas foram utilizadas para processar os dados e identificar automaticamente os comportamentos, dentre elas: Máquina de Vetores de Suporte (SVM), Árvore

de Decisão, *K-means* (*clustering*), Redes Neurais Artificiais (ANN) e Modelos Estatísticos (HMM).

Diosdado *et al.* (2015) comparou o desempenho de quatro algoritmos (Árvore de Decisão, *K-means*, HMM e SVM) e indicou que o algoritmo SVM apresentou os melhores resultados. Porém, considerando o custo computacional, indica o algoritmo de Árvore de Decisão para ser implementado em um sistema de sensores com transmissão dos dados, em função da facilidade de implementação e o custo de processamento.

3.5 Resultados

As métricas de classificação (apresentadas na seção 2.4.7) utilizadas para apresentar os resultados dos comportamentos variam de um trabalho para outro. As principais identificadas foram: Sensibilidade e Precisão ((DIOSDADO *et al.*, 2015) e (GONZÁLEZ *et al.*, 2015)), Acurácia média ((MOREAU *et al.*, 2009), (ROBERT *et al.*, 2009) e (JESUS, 2014)) e ambas ((MARTISKAINEN *et al.*, 2009) e (DUTTA *et al.*, 2015)).

Entre os comportamentos avaliados por Diosdado *et al.* (2015) e Martiskainen *et al.* (2009), o que apresentou melhor resultado de classificação foi o comportamento Pastando, enquanto os com piores resultados foi para distinguir os comportamentos Em Pé e Deitado, utilizando o algoritmo *SVM*. Mesmo resultado apresentado por Jesus (2014), com o algoritmo *END*. Já os trabalhos de Moreau *et al.* (2009) e Robert *et al.* (2009) apresentaram os piores valores na classificação da classe Andando.

Segundo Martiskainen *et al.* (2009), a variação do resultado da classificação entre os diferentes comportamentos, indica que diferentes parâmetros devem ser aplicados ao processo de modelagem do classificador, dependendo das características do comportamento observado. Eles apontam que a qualidade dos dados de aceleração deve ser considerada para melhorar os resultados e que para alguns padrões é necessário métodos de classificação mais sofisticados.

Para Watanabe *et al.* (2005) alguns comportamentos estavam bem distintos enquanto os perfis de aceleração de outros se sobrepunham, assim os autores recomendam, no caso de animais livres na natureza, aferir simultaneamente diferentes parâmetros para melhorar a discriminação dos comportamentos. Os autores sugerem que examinar a sucessão de comportamentos ao longo do tempo, pode ajudar a melhorar a sua identificação. Para o caso de bovinos, por exemplo, pode ser comum o boi beber água após lamber o sal.

3.6 Considerações Finais

Os principais trabalhos que contribuíram para esta pesquisa foram apresentados neste capítulo. Para a definição e produção dos colares, baseou-se nas experiências relatadas, em especial quanto aos sensores utilizados. Os sensores utilizados por eles foram o acelerômetro, o GPS e o giroscópio, este último utilizado apenas no trabalho de [Guo *et al.* \(2009\)](#). O equipamento aqui desenvolvido incluiu estes sensores e ainda adicionou o magnetômetro, um tipo de sensor não utilizado em nenhum dos trabalhos avaliados.

Quanto aos comportamentos aferidos, selecionou-se Pastando, Andando, Em Pé, Deitado e Ruminando, que foram abordados na maior parte das pesquisas, com especial destaque para a necessidade de melhorar a aferição das classes Em Pé e Deitado, que apresentaram conflitos na classificação em alguns trabalhos. Além disso, agregar os dados dos sensores de movimentação do animal com os dados de posicionamento, a fim de identificar comportamentos a partir da combinação, como o caso do animal Bebendo Água que está diretamente relacionado à posição geográfica deste no pasto (área do bebedouro).

Para a anotação das observações em campo, utilizou-se a abordagem de [Jesus \(2014\)](#) e [Dutta *et al.* \(2015\)](#), que fez uso de um *software* para anotar os comportamentos. A ferramenta desenvolvida por [Jesus \(2014\)](#) foi modificada e utilizada nas observações em campo.

Quanto aos algoritmos de classificação, os trabalhos destacam aqueles baseados em Máquina de Vetores de Suporte (SVM) e Árvore de Decisão, em especial [Diosdado *et al.* \(2015\)](#), que aponta o algoritmo SVM como aquele com melhor acurácia, e o baseado em Árvore de Decisão, como o mais fácil de se implementar e com menor custo de processamento. Assim, pretende-se avaliar uma série de algoritmos, com destaque para os destas duas categorias.

Por fim, o interesse no trabalho de [Jesus \(2014\)](#), devido a continuidade das atividades em parceria com a Embrapa e das possibilidades de expansão da proposta apresentada pelo autor.

4 Produção dos Colares

O colar originalmente proposto por [Jesus \(2014\)](#) utiliza um módulo GPS e captura os dados de posicionamento do animal, baseado nos dados do satélite GPS. Um nova versão deste colar foi desenvolvida, incluindo outros sensores, entre eles um acelerômetro, um magnetômetro e um giroscópio, com a função de capturar as movimentações do animal, e um sensor de luminosidade, que apresenta a resistência gerada baseada na intensidade luminosa.

Este capítulo apresenta o processo de produção do colar usado para coleta das informações do gado. As Seções 4.1, 4.2 e 4.3 detalham os sensores utilizados, a placa de circuito confeccionada para interligar os componentes e a confecção dos colares. Na Seção 4.4, apresenta-se a etapa de programação e testes do sistema e, na Seção 4.5, os indicativos da validação do equipamento em campo.

4.1 Sensores Utilizados

Para coleta dos dados de posicionamento do animal foi utilizado um módulo receptor GPS. O modelo utilizado é o Vênus GPS produzido pela SparkFun (Figura 12). Este é um receptor de baixo custo comercial e baixo consumo de energia, com tensão de entrada mínima de 3,3 V e consumo de 2 mA (no estado ocioso) e 6 mA (na taxa máxima de gravação). Ele é baseado no chipset Venus638FLPx, programado por padrão com uma taxa de atualização de 1 Hz (configurável até 20 Hz) e pode ser facilmente integrado a microcontroladores como o Arduino.



Figura 12 – Módulo GPS utilizado - Sparkfun Venus GPS. Disponível em [Electronics \(2015c\)](#).

Para captação dos dados de movimentação do animal foi utilizado o MiniIMU 9DOF v3 (Figura 13), uma placa de medição inercial produzida pela Pololu, com três sensores: acelerômetro, giroscópio e magnetômetro. O chip LSM303D, encapsulado na placa, é responsável pela captura do movimento de aceleração e de medições magnéticas,

enquanto o chip L3GD20H pelos movimentos de rotação, todos eles com 3 eixos de leitura. A placa é acionada com uma tensão de entrada de 2,5 V à 5,5 V e consumo de 6mA.

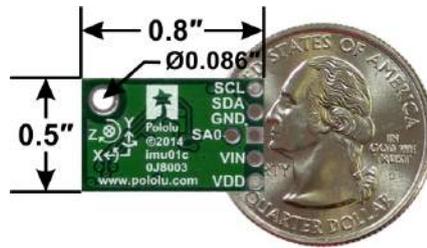


Figura 13 – MiniIMU 9DOF: acelerômetro, giroscópio e magnetômetro. Disponível em [Electronics \(2015a\)](#).

O modelo do acelerômetro suporta configuração com valores de ± 2 , ± 4 , ± 6 , ± 8 , ou $\pm 16g$, o giroscópio variações de ± 245 , ± 500 , ou ± 2000 graus/s e o magnetômetro valores de ± 2 , ± 4 , ± 8 , ou ± 12 gauss.

Para capturar dados da luminosidade do ambiente foi utilizado um LDR (Figura 14).



Figura 14 – Resistor Variável de Luminosidade (LDR). Disponível em [Webtronico \(2015\)](#).

Para o armazenamento dos dados coletados foi utilizado o *data logger* *OpenLog* da SparkFun (Figura 15). Ele permite registrar dados seriais recebidos de 2400 a 115.200 bps e suporta cartões microSD de até 16 GB.



Figura 15 – Módulo OpenLog. Disponível em [Electronics \(2015b\)](#).

Para controlar os sensores e processar as coletas foi utilizado o Arduino, uma plataforma *open-source* de prototipagem eletrônica de *hardware* livre baseada em uma placa micro-controladora. O modelo utilizado foi o Arduino Fio (Figura 16).

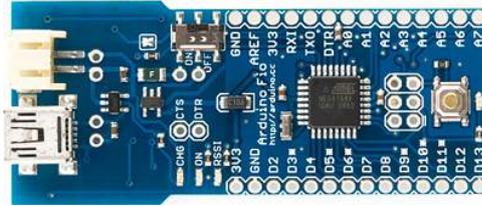


Figura 16 – Arduino Fio. Disponível em [Arduino \(2015\)](#).

4.2 Confeção da Placa de Circuito

O circuito foi projetado e desenhado para que todos os componentes ficassem em uma única placa. O software *Fritzing*¹ foi utilizado para organizar a disposição dos componentes na placa. A Figura 17 mostra a disposição proposta para a placa, utilizando o Arduino Fio como sistema central, unindo o módulo GPS, o módulo dos sensores de movimentação (acelerômetro, giroscópio e magnetômetro em um único componente), o módulo *OpenLog* para gravação dos dados no cartão de memória e um resistor LDR para coleta dos dados de luminosidade. A figura mostra cada uma das ligações necessárias para interligar os componentes no circuito proposto.

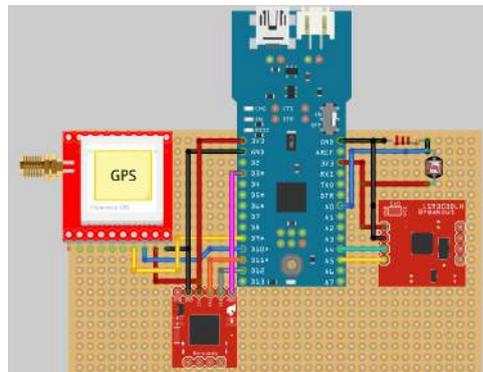


Figura 17 – Projeto do circuito desenvolvido no *Fritzing*.

O circuito foi construído em uma placa de fenolite perfurada, que permite a conexão e solda dos componentes. Na parte superior da placa foram soldados pinos do tipo fêmea, que permitem que os componentes sejam encaixados, possibilitando sua remoção de forma fácil e rápida. Na parte inferior as conexões dos pinos foram interligadas com fios e solda. A Figura 18 apresenta as placas durante o processo de confecção.

A placa construída facilitou o encaixe dos sensores e eliminou os problemas de mal contato, que em geral acontece nos trabalhos realizados em *protoboards*. Para o protótipo essa opção era importante, pois os sensores podiam ser facilmente trocados. Para uma solução comercial recomenda-se que os sensores sejam soldados diretamente à placa do circuito, garantindo que os mesmos não se soltem com o manuseio do equipamento. No

¹ Um *software* livre para modelagem de circuitos eletrônicos. Disponível em: <http://fritzing.org>



Figura 18 – Quatro unidades da placa de circuito durante o processo de produção.

experimento, para evitar que os sensores se soltassem dos encaixes, eles foram fixados aos pinos de encaixe com cola quente.

As ligações do circuito foram feitas através de fios, que foram soldados. O experimento apresentou um problema (relatado na Seção 5.6) nas ligações do circuito do *OpenLog*. Para uma próxima versão pode-se desenhar o circuito e corrobô-lo, construindo uma placa de circuito impresso específica para o sistema, reduzindo as possibilidades de problemas nas ligações. Além disso, a placa Arduino, utilizada como placa microcontroladora do circuito, pertence a um projeto de *hardware* livre², permitindo que a placa seja construída. Assim, a placa do circuito pode já trazer o circuito do Arduino, reduzindo o tamanho final do sistema. Não foi abordado neste trabalho o custo para a produção da própria placa Arduino, mas em geral, a produção desta possui preço menor que os modelos de placas comercializadas, considerando o custo dos componentes necessários para a construção/montagem dela.

Outro problema, foi a programação para interligar todos os componentes no Arduino. O receptor GPS, os sensores de movimentação e o módulo *OpenLog* utilizam a comunicação serial para comunicação com o Arduino. Porém, o Arduino Fio possui apenas um canal de comunicação serial, fazendo com que estes três componentes concorressem pelo mesmo. Foi necessário programar o código de maneira que a comunicação de um, não interferisse na comunicação do outro: o que poderia acontecer é um componente iniciar a gravação no *buffer* enquanto o outro ainda estava utilizando-o.

Nos testes iniciais, com os sensores programados individualmente, este problema não apareceu, surgindo quando eles foram integrados. Apesar dos testes e ajustes, não foi possível garantir que a cada 1 segundo (configuração *default* dos sensores utilizados) um registro dos sensores de movimentação (aqui definido como LAGM - Leitura Acelerômetro, Giroscópio e Magnetômetro) e um registro GPS fossem gravados. Em alguns casos, ao ler os dados do GPS e gravar no *OpenLog*, o registro estava corrompido, ou então dois registros LAGM eram gravados para um registro GPS. Estes problemas foram corrigidos

² O projeto do *hardware* (arquivos de CAD) está disponível sob licença *open-source*, em que pode-se adaptá-la de acordo com as necessidades. Disponível em: <https://www.arduino.cc/en/Main/ArduinoBoardSerialSingleSided3>

no tratamento dos dados (Seção 5.7.1): os registros GPS incompletos foram descartados, assim como o registro LAGM equivalente; os casos com dois ou mais registros LAGM, o último foi unificado ao registro GPS, descartando os demais. Em média, há um registro GPS para o LAGM a cada 3 segundos.

4.3 Confeção dos Colares

Para que o circuito fosse colocado no animal foi utilizado um cinto de couro, com uma caixa acoplada, para guardar o circuito com os componentes eletrônicos. O cinto possui 130 cm x 4 cm x 0,5 cm (comprimento x largura x espessura), a fim de garantir que não arrebentasse caso enroscasse em algo, como um galho ou a cerca.

O circuito foi colocado em uma caixa feita com material plástico, com dimensões de 14 cm x 5 cm x 10 cm (largura x altura x profundidade) e uma tampa removível. Para prender o circuito à caixa, ele foi fixado na tampa, com fios e cola quente (Figura 19).



Figura 19 – Placa de circuito fixada na tampa da caixa.

A caixa foi presa ao cinto com fita de uso geral. A fita utilizada é constituída de um filme de polietileno e trama de tecido com adesivo à base de resina e borracha (*silver tape*), que também serviu para proteger a caixa contra eventuais impactos e sujeiras. A Figura 20 mostra o cinto com a caixa.

O sensor LDR não ficou dentro da caixa. Dois furos foram feitos no cinto, em uma das extremidades, onde o sensor foi encaixado e soldado a dois fios que foram interligados ao circuito. A antena do GPS também ficou fora da caixa e foi fixada ao lado do sensor LDR. A Figura 21 mostra a antena e o LDR fixos ao colar. A posição utilizada buscou garantir que os sensores ficassem na parte superior do animal, capturando os dados de luminosidade e o sinal dos satélites GPS. A antena foi fixada com fita isolante e toda a fiação envolvida com a fita de uso geral, para que não sofressem danos.

Quatro unidades do colar foram confeccionadas. Para identificá-las cada uma foi codificada com uma letra e pintadas em cores diferentes. Desta maneira é possível diferenciar os colares quando utilizados simultaneamente. Neste trabalho apenas dois colares



Figura 20 – Cinto com a caixa do circuito.

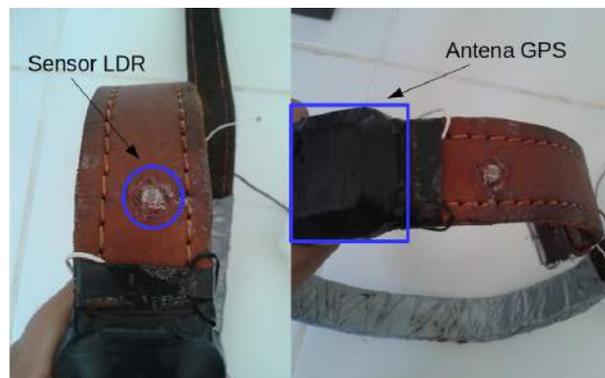


Figura 21 – Sensor LDR e Antena GPS fixados ao colar.

foram utilizados simultaneamente, logo utilizou-se apenas duas cores. Os colares identificados como "A" e "C" ficaram na cor marrom (cor original do cinto) e os colares "B" e "D" pintados na cor azul.

4.4 Implementação e Testes

O circuito e os componentes foram testados, para validar o funcionamento individual de cada um e a integração de todos em um único sistema. Inicialmente foram realizados testes de bancada e na sequência testes de campo.

4.4.1 Circuito e Sensores

Inicialmente, para ativar os sensores, foi utilizado o código exemplo fornecido pelos fabricantes dos componentes. Em geral, junto ao anúncio dos componentes são apresentadas características de funcionamento e os *links* para a documentação e código exemplo. Os testes foram feitos em uma *protoboard* (Figura 22) e cada unidade dos componentes

testados. As placas de circuito também foram testadas, para garantir que todas as ligações estavam corretas.

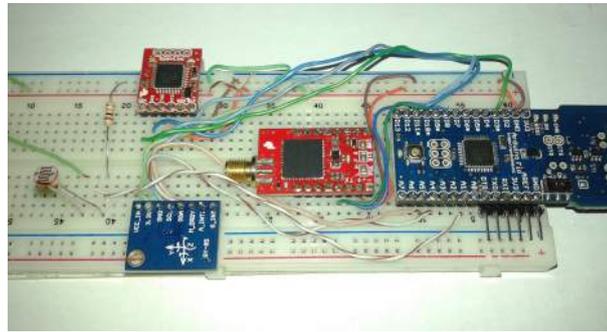


Figura 22 – Circuito montado na protoboard para testes de bancada.

Após os testes, o código de cada sensor foi personalizado, para que os dados de interesse fossem enviados ao *OpenLog*. Um único código foi desenvolvido, unificando a coleta dos dados dos sensores e enviando-os ao *OpenLog*, para gravação em um micro cartão SD. Esta validação foi realizada utilizando todos os sensores conectados na placa de circuito (Figura 23).



Figura 23 – Placa de circuito com os sensores encaixados.

4.4.2 Bateria

Para alimentar o circuito de cada colar foi utilizado uma bateria de polímero de Lítio de 6Ah, com 3 células (cada célula fornece 3,7V nominais a 2000mAh). Esta bateria é muito fina e extremamente leve, além de ser recarregável.

As baterias utilizadas para alimentar o circuito foram testadas antes do experimento, para identificar o tempo médio de funcionamento do circuito, sem a necessidade de recarga. Antes de qualquer teste, que envolvesse a autonomia da bateria, a tensão fornecida por ela era medida com auxílio de um multímetro, garantindo que estava fornecendo

a tensão nominal. O teste era realizado até o Arduino ser desligado automaticamente, que ocorre quando a tensão fornecida for abaixo de um valor limite. Este recurso do Arduino aumenta a vida útil da bateria, já que ela não é utilizada até descarregar totalmente.

Os teste foram realizados em uma bancada, dentro de uma sala. As coletas eram feitas a cada 1 segundo, para todos os sensores e os resultados gravados no cartão SD e exibidos na tela por meio do monitor serial (recurso disponível na plataforma Arduino).

Quatro colares e quatro baterias foram testados, e o tempo médio de autonomia do sistema foi de 60 horas: o maior registro com 80 horas e o menor com 50 horas. Vale ressaltar que nestes testes o código executado possuía uma função para também exibir os dados na tela e os colares estavam em um ambiente fechado, fazendo com que o receptor GPS consumisse mais energia na busca pelo satélite, se comparado ao uso em campo aberto, quando não há obstáculos para comunicação com o satélite.

4.5 Indicadores da Experiência em Campo

Após os testes de bancada o circuito foi colocado na caixa e acoplado ao colar, para validação em campo. Dois períodos de testes - de 25 a 30 de março e 06 a 09 de abril de 2015 - foram realizados com os quatro colares na área de pasto do Mangueiro Digital na Embrapa. A área utilizada possui 4,09 ha e está identificada como Área 1 na Figura 24.



Figura 24 – Área do Mangueiro Digital.

Nos dois períodos, os testes buscaram validar o funcionamento do sistema e serviram para ajustar detalhes no encapsulamento da caixa, do posicionamento do colar no animal e o tempo de vida útil da bateria. Neste teste não houve nenhum tipo de análise dos comportamentos dos animais.

4.5.1 Material Utilizado

Os resultados dos testes iniciais serviram para melhorar o encapsulamento do circuito. Na primeira versão foram utilizadas caixas de formatos diferentes em cada colar. Além disso, a maneira com que os componentes foram alocados nela não foi adequado. A proposta era deixar o conector da antena GPS para fora da caixa, possibilitando que a antena fosse retirada sem a necessidade de abertura da caixa. Porém, nos testes em campo, o conector para fora sofreu pressão nos momentos em que o animal encostava o pescoço no chão e no momento da vedação da caixa, o que fez com que o módulo GPS se desligasse da placa de circuito. A solução foi deixar a conexão na parte interna da caixa, fazendo um furo na caixa para passar o fio da antena. A Figura 25 compara a primeira e a segunda versão da caixa desenvolvida, destacando a posição do encaixe da antena GPS.



Figura 25 – Duas versões da caixa para encapsulamento do circuito.

A caixa do colar é na verdade um pote plástico com tampa, que foi adaptado ao experimento. Não há registros de problemas com o seu uso, porém para uma nova versão (ou uma versão comercial) sugere-se a construção de uma caixa especificamente modelada para este fim.

A cada troca do colar, no momento em que o cartão de memória era retirado e a bateria recarregada, era necessário tirar a fita *silver tape* e posteriormente recolocá-la. Uma alternativa para a fita é desenvolver uma versão da caixa que seja fixa ao cinto sem a fita (com parafusos ou similares). Este ajuste, além de melhorar a apresentação do colar, possibilita a abertura da caixa de maneira prática e com menor custo, já que a fita não pode ser reutilizada.

Outro cuidado, é garantir que a caixa não seja danificada em possíveis impactos. Uma alternativa é revesti-la com material emborrachado. Além disso, é necessário vedar a caixa, para evitar problemas com umidade, água e outros elementos que possam danificar o circuito. No período de experimento há registros de momentos em que o colar ficou exposto a chuva, porém sem nenhum tipo de dano gerado no circuito, já que a fita vedou a caixa.

Outros possíveis ajustes na caixa são: um encaixe para a placa de circuito, de maneira que seja fixada com parafusos (ou similares) e não com cola, já que a cola precisou ser repassada após alguns experimentos em campo (necessidade de manutenção frequente); uma conexão externa que permita recarregar as baterias sem a abertura da caixa; e uma chave on/off que possibilite que o sistema seja ligado/desligado também sem a abertura da caixa. Esta padronização da caixa, com os encaixes para a placa de circuito e pontos para prendê-la no cinto, garantem que a placa do circuito e a caixa estejam sempre na mesma posição, mesmo em unidades diferentes do colar, ou seja, os eixos dos sensores de movimentação estarão sempre na mesma posição.

4.5.2 Posição do Colar no Animal

O tamanho (diâmetro) do pescoço do animal impactou na posição da antena e do sensor LDR. Estes dois elementos foram fixados no colar em uma posição, que ao fechar o cinto, deveriam ficar em lado oposto a caixa. A Figura 26 representa a ideia da posição ideal.

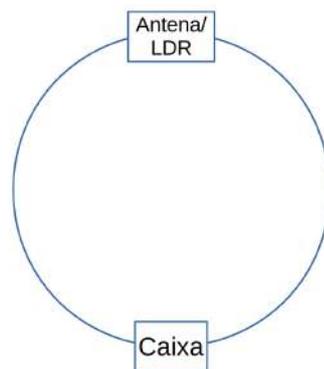


Figura 26 – Posição ideal da antena GPS e LDR com relação a caixa com o circuito.

O peso da caixa (em função do circuito e da bateria) fez com ela ficasse sempre abaixo do pescoço do animal, enquanto a antena GPS e LDR na parte de cima do pescoço. Porém, como cada animal possui um diâmetro de pescoço, a Antena/LDR nem sempre ficava posicionada exatamente na parte superior. Em alguns casos elas ficaram na lateral, já que a posição está diretamente relacionada ao ponto que o cinto foi amarrado (furo do colar em que a fivela era presa). De acordo com o diâmetro do pescoço uma posição era utilizada.

Além da influência na posição da antena/LDR, há o fato de que em um animal, com pescoço menor, o cinto ficou largo, permitindo que o mesmo pudesse se deslocar. Em um dos momentos de observação, ocorreu do animal passar o pescoço na cerca e o cinto girar, deixando a caixa na parte superior do pescoço. Após alguns movimentos, o peso da caixa fez com que ela voltasse à posição inicial (abaixo do pescoço). Para este experimento não houve interferência, já que o observador não anotou os comportamentos

neste momento, porém para utilização em situações em que não há observação, os dados coletados estarão fora do padrão.

Uma solução seria utilizar um outro fecho para prender o colar, no lugar da fivela, que prenda o cinto e deixe-o justo ao pescoço. O cuidado deve ser para que o material não abrase o animal e que seja firme suficiente para evitar que se rompa caso o animal enrosque em algum objeto. Esta alternativa resolve o problema do deslocamento do colar no pescoço do animal, porém não resolve a posição da antena e do LDR, pois estes são fixos no colar e variam conforme a posição que o cinto é fechado. Apesar desta variação, para o GPS não há problemas, já que não há registros de perda de sinal do satélite durante o experimento, porém para a luminosidade há impacto, pois se o sensor ficar na lateral do pescoço, o corpo do animal projetará sombra sobre o sensor.

A maneira com que o colar é colocado no animal também influencia a posição dos eixos dos sensores de movimentação. Assim, um novo cinto pode indicar a posição correta, para evitar que sejam instalados em posições diferentes. Em uma das coletas realizadas, o colar foi colocado de maneira invertida, ou seja, os eixos dos sensores não ficaram de acordo com o previsto. O colar foi amarrado do lado direito do animal, quando deveria ficar do lado esquerdo (como será apresentado na Seção 5.2). A Figura 27 mostra o animal com o colar amarrado na posição errada.



Figura 27 – Animal com o colar em posição invertida: amarrado no lado direito.

4.6 Considerações Finais

Este capítulo apresentou os materiais e equipamentos utilizados na produção dos colares, detalhando as características dos sensores e as especificações dos componentes. O procedimento de programação dos componentes e dos colares também foram detalhados, mostrando os resultados dos testes de bancada e de campo. As sugestões podem

ser aplicadas em um versão futura do equipamento, a fim de evitar os problemas aqui relatados.

5 Experimento em Campo

Com os colares prontos e testados, foi realizada a coleta de dados na área de pasto da Embrapa. Neste capítulo, são relatados os resultados da experiência em campo. Nas Seções 5.1, 5.2 e 5.3, são apresentadas as características dos animais, o posicionamento do colar nos animais e os comportamentos observados. As Seções 5.4 e 5.5 trazem os períodos de observação e os detalhes do *software* utilizado para anotação das observações. Por fim, nas Seções 5.6 e 5.7, são mostrados os procedimentos utilizados para unificar os dados coletados e para prepará-los para a aferição dos comportamentos.

5.1 Animais Utilizados no Experimento

A área e os animais observados e equipados com os colares são do Mangueiro Digital, um laboratório voltado para a Pecuária de Precisão da Embrapa Gado de Corte (área apresentada na Figura 24). Neste espaço há dois mangueiros para o manejo dos animais.

Oito animais estavam no pasto no período do experimento e foram mantidos apascentados em pastagem cultivada e com suplementação mineral adequada, água à vontade e tratamento sanitário, pelos profissionais da Embrapa. A Tabela 3 apresenta as características dos três animais que utilizaram o colar durante o experimento.

Tabela 3 – Características dos animais utilizados no experimento.

ID	Raça	Pelagem	Categoria	Sexo	Nascimento
6998	Caracu	Vermelha	Novilha	Fêmea	19/09/2013
7014	Caracu	Vermelha	Novilha	Fêmea	23/09/2013
7135	Caracu	Baia	Novilha	Fêmea	02/11/2013

5.2 Posição do Colar no Animal

Os sensores de movimentação utilizados (acelerômetro, giroscópio e magnetômetro) possuem medições em 3 eixos (X, Y e Z). A posição em que a caixa é fixada no colar interfere na posição destes eixos. A Figura 28 mostra a posição dos eixos do acelerômetro no momento em que o colar foi colocado no animal e o mesmo está com o pescoço reto.

Para padronizar a posição dos eixos em todas as coletas, foi necessário ajustar a forma com que a caixa era fixada no colar. Em todas as coletas a caixa foi fixada com o lado da bateria virada para o mesmo lado da ponta do cinto (lado oposto à fivela). A Figura 29 apresenta a caixa aberta, mostrando a organização interna e a posição da bateria dentro dela, posição que deve ser combinada com a ponta do cinto.

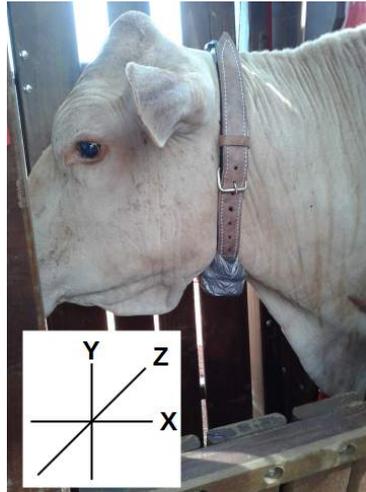


Figura 28 – Animal com o pescoço reto. Destaque para a posição de cada eixo do acelerômetro.



Figura 29 – Caixa com o circuito aberto, com destaque para a posição da bateria.

A posição que o colar é amarrado no pescoço do animal também influencia a posição dos eixos. Para padronizar a posição determinou-se que o lado que o cinto é amarrado (preso com a fivela) deve ficar do lado esquerdo do animal. O lado utilizado poderia ser o outro, porém a título de padronização da posição dos eixos dos sensores de movimentação, foi adotado o lado esquerdo. Na Figura 28 é possível observar este detalhe.

5.3 Comportamentos Observados

Os principais comportamentos bovinos apontados na literatura são os comportamentos Pastando/Procurando, Andando, Em Pé e Deitado. Baseado no trabalho de Jesus (2014), estes foram os comportamentos inicialmente definidos para observação durante o experimento e aqui são definidos como comportamentos primários. A seguir são apresentadas as características observadas em cada um destes comportamentos.

- Pastando/Procurando: caracterizado pelo animal sobre as quatro patas, com a cabeça baixa procurando ou mastigando o capim. O animal pode ou não estar em movimento, já que ele pode estar se deslocando à procura de capim;
- Andando: o animal também está sobre as quatro patas, porém com o pescoço reto (apontando o fucinho para frente) e se deslocando pela área de pasto;
- Em Pé: o animal está sobre as quatro patas, com a cabeça erguida e não há deslocamento;
- Deitado: o animal está com as patas abaixadas e com a barriga tocando o solo.

A Figura 30 apresenta o animal na posição em cada um dos quatro comportamentos primários.

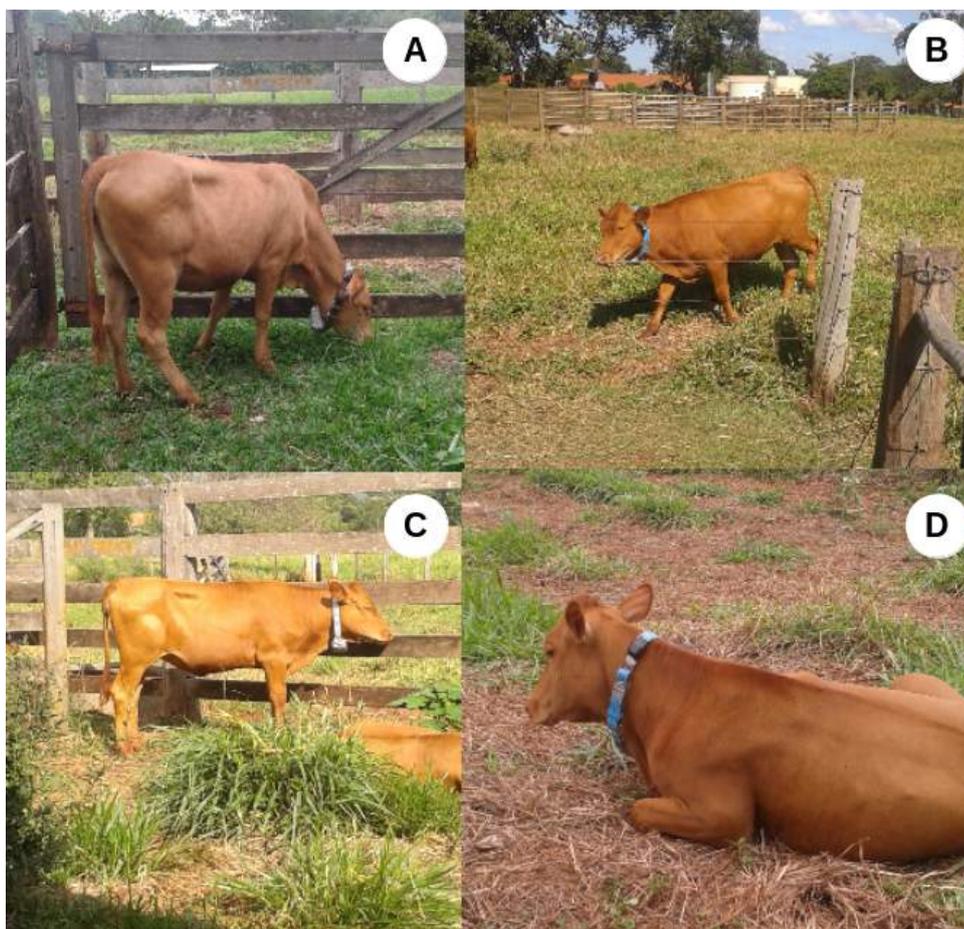


Figura 30 – Animal nas posições dos quatro comportamentos primários: (A) Pastando/Procurando, (B) Andando, (C) Em Pé e (D) Deitado.

Após contato com os pesquisadores da Embrapa Gado de Corte, mais dois comportamentos foram adicionados para observação: Ruminando Em Pé e Rumimando Deitado. As posições dos comportamentos são semelhantes ao Deitado e Em Pé, com a diferença

que no comportamento Ruminando o animal está movimentando a mandíbula (movimento de mastigação).

Além destes comportamentos, um sétimo comportamento foi adicionado para observação: Bebendo Água. Os momentos em que o animal foi até o bebedouro para beber água também foram anotados. Este comportamento é caracterizado pelo animal de cabeça baixa e parado no local dos bebedouros. Como o colar coleta os dados de posicionamento do animal é possível relacioná-lo à posição georreferenciada do bebedouro, identificando comportamentos a partir da combinação dos dados de posicionamento e de movimentação.

O observador anotou os comportamentos em um intervalo máximo de 5 minutos, porém não limitado a ele (há registros em período menores). Os valores lidos pelos eixos dos sensores variam mesmo em pequenos movimentos, portanto se dentro do intervalo de 5 minutos o observador notasse variação na posição do animal um novo registro era feito, mesmo quando o comportamento não era alterado. Quando o intervalo expirava, uma série de 10 anotações era feita, uma a cada segundo. A série com 10 anotações foi planejada considerando que os registros do colar são gravados em média a cada 3 segundos, o que garantiria minimamente 3 anotações do observador com registro equivalente no colar, a cada série. Ressalta-se que as anotações não garantem que o comportamento começou ou terminou em um determinado momento, pois o interesse estava em coletar os dados durante o comportamento e não registrar o início e fim do mesmo.

As anotações foram realizadas com o *software* Ápis, desenvolvido por Jesus (2014) e modificado durante este trabalho. As características da ferramenta são apresentadas na Seção 5.5.

5.4 Período de Observação

Os animais foram equipados com os colares e observados em dois períodos diferentes. O primeiro aconteceu entre os dias 14 e 28 de maio/2015. Dois animais foram selecionados aleatoriamente no rebanho e equipados com o colar, de maneira que um ficasse com o colar na cor azul e o outro com o colar na cor marrom. Para este período foram programadas quatro datas para realizar a troca dos colares, agendadas baseado na autonomia das baterias. Nas datas programadas, os colares eram substituídos por outros dois colares: o animal que estava com o colar na cor azul, era equipado com o outro colar azul e o animal com o colar marrom, equipado com o outro marrom. Os animais equipados neste período foram os marcados com os números 6998 e 7014. O código do animal equivale ao número que o identifica e que está marcado na sua pele.

Apesar das mesmas cores, cada colar possuía um código de identificação. Em cada troca, o conjunto de dados era identificado com o código do colar e o número sequencial da coleta realizada com o colar: por exemplo, no conjunto rotulado com o código A2, o

número 2 identifica a segunda coleta realizada com o colar A. A Figura 31 mostra um dos momentos de troca dos colares realizado com auxílio da equipe da Embrapa.



Figura 31 – Troca de colar realizada no mangueiro com auxílio dos profissionais da Embrapa.

O segundo período de coleta ocorreu entre os dias 25 de junho e 06 de julho/2015. O intervalo foi planejado considerando que possíveis problemas apresentados na primeira etapa poderiam ser corrigidos a tempo de uma nova coleta e de acordo com o agendamento da Embrapa para manejo dos animais. Vale ressaltar que não foi necessário utilizar o intervalo entre os períodos para ajustes.

Os colares e animais foram organizados na mesma dinâmica do primeiro período, com três datas programadas para a troca dos colares. Nesta ocasião, um dos animais era facilmente diferenciado por apresentar pelagem mais clara. Os animais que utilizaram o colar neste período estavam marcados com os números 6998 e 7135. Pode-se observar que o animal com o número 6998 participou dos dois períodos de coleta, porém de maneira aleatória, já que os animais escolhidos eram aqueles que primeiro entravam no tronco.

Os dois períodos de observação do experimento somam 513 horas de coleta de dados dos colares e mais de 70 horas de observação. A Tabela 4 mostra os valores por período de coleta, dividindo os resultados das coletas no colar e nas observações com o Ápis.

Tabela 4 – Volume de Dados Coletados em Dias/Horas

	Dados dos Colares		Dados de Observação	
	Dias	Horas	Dias	Horas
1º Período	14	279	9	42
2º Período	11	234	7	29
Total	25	513	16	71

As anotações foram feitas baseadas na observação visual realizada na área do experimento. O observador, em alguns momentos, interferiu no comportamento do animal: por exemplo, ao ficar próximo a cerca do pasto, o animal parava e ficando olhando o observador. Como a proposta deste trabalho foi capturar a característica de cada comportamento (posição e movimentos durante o comportamento), esta interferência não compromete o resultado.

Inicialmente propôs-se realizar a observação simultânea de quatro animais mas, baseado na experiência dos profissionais da Embrapa, seria uma atividade difícil de ser realizada. Assim, definiu-se que apenas dois animais seriam observados simultaneamente. Esta decisão também fez com que os dados fossem coletados de maneira contínua, já que ao tirar os colares outros dois já estavam disponíveis para utilização. Caso os quatro fossem utilizados, haveria um tempo entre tirar os colares e recarregar as baterias, pois apenas quatro baterias estavam disponíveis.

Mesmo com apenas dois animais, em alguns momentos era difícil diferenciá-los. Apesar dos colares possuírem cores diferentes, houve momentos em que eles estavam afastados e a sua identificação era bloqueada por outro animal ou por algum obstáculo que estava na frente. Houve também situações em que os dois animais estavam em áreas opostas no pasto, impossibilitando que os dois fossem observados. Nestes casos apenas os dados de um animal era coletado.

Nas primeiras 24 horas, dos dois períodos de coleta, não foram realizadas coletas e observações, a fim de garantir que o animal se acostumasse com o equipamento. A Tabela 5 apresenta os períodos de coletas de dados, apresentando o código do colar, o número sequencial, o período e o código do animal monitorado.

Tabela 5 – Detalhes das coletas e observações organizados por colar.

Cod.Colar	Nº Sequência	Período	Cod.Animal
A	1	18 à 22/05/15	6998
A	2	25 à 27/05/15	6998
A	3	29 à 02/07/15	6998
B	1	18 à 22/05/15	7014
B	2	25 à 29/05/15	7014
B	3	29 à 02/07/15	7135
C	1	14 à 18/05/15	6998
C	2	22 à 25/05/15	6998
C	3	25 à 29/06/15	6998
C	4	02 à 06/07/15	6998
D	1	14 à 18/05/15	7014
D	2	22 à 25/05/15	7014
D	3	25 à 27/06/15	7135
D	4	02 à 06/07/15	7135

Nas 14 coletas realizadas, a autonomia média da bateria foi de 78 horas, com a maior coleta registrando 95 horas e a menor 72 horas. Em apenas 5 coletas a bateria esgotou antes da retirada do colar, ou seja, em 9 coletas quando o colar foi retirado do animal ainda havia carga na bateria. Ressalta-se que nestes experimentos há um registro válido a cada 3 segundos, em média, apesar dos sensores estarem configurados para realizar coletas a cada 1 segundo.

5.5 Software para Anotação dos Comportamentos

O *software* Ápis foi desenvolvido por Jesus (2014) para anotação do comportamento animal. Ele foi desenvolvido para ser utilizado em qualquer dispositivo móvel - *smartphone* ou *tablet* - que utilize o sistema operacional *Android*. Essa versão inicial permite que um animal seja observado utilizando quatro comportamentos predefinidos, porém neste trabalho houve a necessidade de observação de mais de um animal e de mais comportamentos.

As alterações no *software* Ápis trouxeram quatro novas funcionalidades: (1) escolha da quantidade de animais observados; (2) escolha da quantidade de comportamentos; (3) personalização dos comportamentos observados; (4) despertador para avisar o observador que os comportamentos devem ser anotados. Os detalhes das novas funcionalidades são apresentadas a seguir.

Para orientar a comparação das novas funcionalidades com a versão original desenvolvida por Jesus (2014), apresenta-se na Figura 32 a tela da versão original do Ápis.

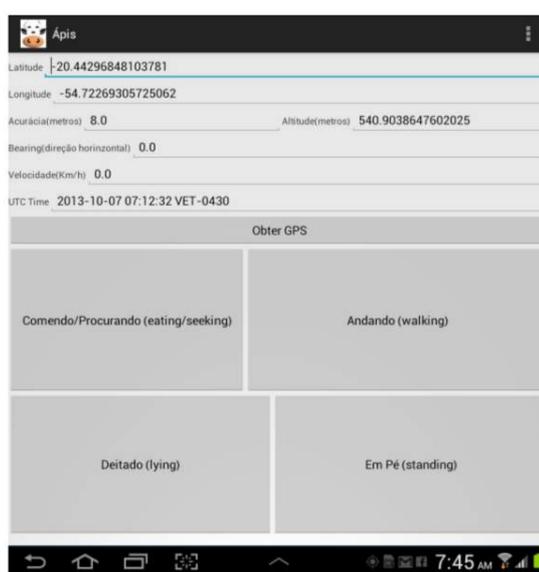


Figura 32 – Tela do Ápis na versão desenvolvida por Jesus (2014).

O contato com os pesquisadores da Embrapa foi fundamental para a definição destas funcionalidades, já que foram apresentadas outras demandas em que o *software* pode ser utilizado. As novas funcionalidades, além de atender este trabalho, permitirá que outros trabalhos utilizem-a, já que a personalização das observações e da quantidade de animais possibilita sua utilização em diferentes demandas. A seguir são apresentadas as funcionalidades implementadas:

- (1) Na versão anterior não havia o controle de qual animal estava em observação, assim o usuário apenas selecionava os comportamentos. Neste trabalho, com a necessidade

de observar simultaneamente animais diferentes, foi implementado a opção para observar mais de um animal. Para diferenciar os animais, cada um deles é identificado com um código e uma cor. A quantidade de animais é definida na tela inicial do aplicativo (Figura 33).

Uma tela inicial de configuração foi incluída, permitindo que o usuário selecione a quantidade de animais observados e quantos/quais os comportamentos observados. Inicialmente, o usuário seleciona o número de animais observados e quais comportamentos deseja observar.



Figura 33 – Tela inicial para personalização da quantidade de animais e comportamentos observados.

A Figura 34 mostra a tela com a opção para escolha de dois animais, onde cada animal está identificado com um código e uma cor.

Após a configuração inicial, uma tela é aberta com um botão para cada animal selecionado. Cada botão possui uma cor diferente e possui um rótulo identificando-o. Ao pressionar o botão do animal escolhido, uma nova tela é aberta, apresentando os comportamentos, para que o usuário então selecione-o. Ao selecionar um dos comportamentos o aplicativo volta para a tela anterior, apresentando novamente a tela com os animais disponíveis.

- (2) Na primeira versão o *software* permitia a anotação de quatro comportamentos: Pastando/Procurando; Andando; Deitado; Em Pé. A nova funcionalidade trouxe a possibilidade do usuário escolher quantos e quais comportamentos observar. Oito comportamentos são predefinidos no aplicativo, com a opção do usuário selecionar quais deseja utilizar (Figura 33).



Figura 34 – Tela com os botões para seleção dos animais.

- (3) Além de escolher a quantidade de comportamentos, o *software* permite que os rótulos (descrição de cada comportamento) sejam personalizados. O rótulo apresentado na tela é o mesmo gravado no arquivo de saída. A Figura 35 mostra a tela com a opção de escolha dos comportamentos, onde cada botão apresenta o rótulo informado pelo usuário na tela de configuração.



Figura 35 – Tela com os botões para seleção dos comportamentos.

- (4) A cada cinco minutos, um sinal sonoro e visual, informa ao usuário que os comportamentos dos animais devem ser anotados.

O uso do Ápis facilitou o trabalho do observador, já que os dados não precisaram ser transcritos, pois ele disponibiliza um arquivo de texto com as informações coletadas. Ao selecionar um animal e um comportamento, o aplicativo grava em um arquivo texto a data e horário (do satélite GPS), a cor do animal selecionado e o comportamento observado.

5.6 Combinação dos Dados Coletados

Apesar dos testes realizados antes do experimento em campo, alguns problemas aconteceram nos colares. Dos 14 conjuntos coletados, três apresentaram problemas: dois conjuntos apresentaram problema no cartão que armazena os dados (C1 e C2) e um conjunto na bateria que alimenta o circuito (B1).

Um dos problemas foi a queima do cartão SD que armazena os dados coletados. O problema também aconteceu nos primeiros testes realizados em campo, antes do período de observação: os dois casos ocorreram em um mesmo circuito (mesmo colar). As soldas do circuito, que conectam o *OpenLog* ao Arduino, foram refeitas e não houve novos registros deste problema.

O outro problema foi um conjunto que não gravou nenhuma informação no cartão. Ao abrir a caixa para a retirada dele observou-se que o *OpenLog* estava mal conectado na placa de circuito. Possivelmente o problema aconteceu quando a fita foi passada para vedar a caixa, já que ela fazia pressão sobre a caixa. Este problema também foi observado no período inicial de testes. A solução adotada após o último incidente, foi fixar os sensores nos encaixes utilizando cola quente, garantindo que os componentes não se soltassem durante a vedação da caixa.

No terceiro conjunto que não gravou os dados, o problema foi a bateria que alimentava o circuito. Um pino, presente na placa Arduino, furou a bateria no momento em que a fita foi passada para vedar o circuito, também devido à pressão realizada. Esta constatação foi feita pois haviam alguns registros gravados no cartão, indicando que o mesmo foi desativado instantes após a vedação. Apesar da perfuração de uma das células da bateria, foi possível reutilizá-la nas coletas seguintes. O pino que furou a bateria foi isolado para as outras coletas.

O conjunto A1 também não foi considerado nos resultados apresentados na sequência do trabalho, já que o colar foi colocado na posição errada no animal (conforme relatado na Seção 4.5.2).

Em todo o período foram registradas mais de 33 mil anotações de comportamentos observados. A quantidade de anotações coletadas em cada colar, é apresentada na Tabela 6, detalhando a quantidade de cada comportamento.

Destes registros, pouco mais de 13 mil foram efetivamente utilizados, porque nem

Tabela 6 – Número de anotações realizadas por comportamento em cada colar.

Colar	Andando	Pastando/Procurando	Deitado		Em Pé		Total
			Ruminando	Parado	Ruminando	Parado	
A	881	3327	1690	945	724	811	8378
B	491	3638	564	1122	376	882	7073
C	354	1659	1729	789	469	636	5936
D	1113	5695	1458	1752	531	1668	12217
Total	2839	14619	5441	4608	2100	3997	33604

todas as anotações tinham registro equivalente no colar. Ou seja, para um registro de observação não há registro no colar, comparando a data/horário da anotação. Isso aconteceu pelo fato dos registros do colar serem gravados, em média, a cada 3 segundos. A Tabela 7 mostra a quantidade de registros de cada comportamento que possuem combinação com os dados dos colares.

Tabela 7 – Número de registros das observações com combinações nos colares.

Colar	Andando	Pastando/Procurando	Deitado		Em Pé		Total
			Ruminando	Parado	Ruminando	Parado	
A	359	1198	642	348	288	310	3145
B	207	1454	235	461	155	354	2866
C	165	753	630	282	184	244	2258
D	498	2327	527	611	205	651	4819
Total	1229	5732	2034	1702	832	1559	13088

5.7 Tratamento dos Dados Coletados

Nesta seção é apresentada a estrutura dos registros gravados a partir dos dados coletados nos colares, no *software* de observação e como estes foram organizados para utilização na classificação dos comportamentos dos animais.

5.7.1 Dados do Colar

Cada colar registrou os dados do GPS, dos sensores de movimentação (acelerômetro, giroscópio e magnetômetro) e do sensor de luminosidade. Os dados foram gravados em um cartão SD utilizando o *OpenLog*, organizado em dois registros: registro GPS e registro dos demais sensores (aqui definido como LAGM).

As saídas do GPS seguem o padrão NMEA-0183, um conjunto de especificações de dados para comunicação de dispositivos eletrônicos de navegação. A saída do receptor possui cinco *strings*, que são descritas sucintamente a seguir ([TECHNOLOGY, 2008](#)):

- \$GPGGA (solução de posição determinada): fornece a posição representada por latitude e longitude, o horário em que a posição foi obtida e sua validade;
- \$GPGSV (satélites visíveis): fornece dados de identificação dos satélites utilizados;
- \$GPGSA (satélites ativos): fornece informações sobre a qualidade da posição obtida;

- \$GPRMC (especificações mínimas): fornece a posição representada pela latitude e longitude, informações de hora e data, declinação magnética;
- \$GPVTG (rumo percorrido): fornece dados sobre o deslocamento, como a velocidade em km/h.

O outro registro, com os dados dos demais sensores, foi identificado com a *tag* \$LAGM no início dos registros. O registro é formado pelo código do colar, valor do LDR, eixo X acelerômetro (AX), eixo Y acelerômetro (AY), eixo Z acelerômetro (AZ), eixo X magnetômetro (MX), eixo Y magnetômetro (MY), eixo Z magnetômetro (MZ), eixo X giroscópio (GX), eixo Y giroscópio (GY) e eixo Z giroscópio (GZ). A Tabela 8 mostra um exemplo da estrutura do registro.

Tabela 8 – Estrutura dos registros LAGM com a identificação do colar e valores dos sensores LDR e de movimentação.

Tag	Cod.Colar	LDR	AX	AY	AZ	MX	MY	MZ	GX	GY	GZ
\$LAGM	A1	128	-3814	10184	-11577	-1341	559	-635	395	252	203

No arquivo de coleta dos colares, há a cada 1 segundo, um registro LAGM e um conjunto de registros GPS, como mostra a Tabela 9.

Tabela 9 – Exemplo dos dados originais do colar: registros LAGM + GPS.

```

$LAGM,A1,121,-8071,8549,-12355,-1306,375,-510,-659,-2771,-7409,MGAL
$GPGGA,131856.203,2026.5936,S,05443.3858,W,1,10,0.8,527.4,M,4.8,M,,0000*64
$GPGSA,A,3,21,22,26,18,25,15,29,16,27,31,,1.4,0.8,1.1*39
$GPGSV,3,1,12,18,79,150,39,22,61,316,33,26,54,276,31,21,38,147,44*7C
$GPGSV,3,2,12,16,36,235,36,29,29,072,45,27,16,219,30,15,14,118,29*75
$GPGSV,3,3,12,14,13,006,,25,07,019,39,31,05,325,25,08,00,000,25*72
$GPRMC,131856.203,A,2026.5936,S,05443.3858,W,001.0,157.7,250515,,A*64
$GPVTG,157.7,T,,M,001.0,N,001.8,K,A*01

```

Dos dados gravados no colar, aqueles de interesse para a identificação dos comportamentos são: os dados dos 9 eixos (X, Y e Z dos 3 sensores de movimentação), data e horário do satélite GPS, as coordenadas geográficas (latitude e longitude) e a velocidade de deslocamento. Estes dados foram gravados em um único registro e adicionados em um novo arquivo. A Tabela 10 mostra um exemplo do novo registro gerado a partir dos dados originais do colar.

Tabela 10 – Exemplo de um registro do colar após tratamento.

Conj	AX	AY	AZ	MX	MY	MZ	GX	GY	GZ	HORÁRIO	DATA	LAT	LONG	KM/H
A1	-8071	8549	-12355	-1306	375	-510	-659	-2771	-7409	131856.203	250515	2026.5936	05443.3858	001.8

5.7.2 Dados do Software de Observação

Os dados de interesse dos colares são aqueles que possuem registros combinados com as anotações no *software* de observação, baseado na data e o horário dos registros.

Tabela 11 – Registros gerados pelo software de observação.

Data	Horário	Padrão UTC	Animal/Colar	Comportamento
2015-05-25	18:04:51	Horário de Greenwich+0000	A2	Deitado
2015-05-25	18:05:15	Horário de Greenwich+0000	B2	Andando

A Tabela 11 apresenta exemplos dos registros gerados pelo Ápis.

Em cada registro foi alterado a máscara de apresentação da data e horário e eliminado o campo Padrão UTC (este campo confirma o fuso horário configurado no dispositivo móvel). A Tabela 12 apresenta exemplos dos registros do *software* após o tratamento.

Tabela 12 – Registros do software de observação após tratamento.

Data	Horário	ID Colar	Comportamento
250515	180451	A2	Deitado
250515	180515	B2	Andando

5.7.3 Unificação dos Registros

Os dados dos colares e das observações foram combinados, para formar registros únicos. A combinação foi realizada utilizando as datas e horários dos registros, já que em ambos estes dados foram coletados do satélite GPS. Para cada conjunto de dados (colar + observação) foi gerado um arquivo de saída no formato ARFF, formatado e com o cabeçalho para entrada na ferramenta WEKA. A Tabela 13 mostra um fragmento de um dos arquivos gerados, para o conjunto A3.

Tabela 13 – Arquivo com os registros unificados e com o cabeçalho formatado para a ferramenta WEKA.

```

@relation comportamentos

@attribute ID String
@attribute AX numeric
@attribute AY numeric
@attribute AZ numeric
@attribute MX numeric
@attribute MY numeric
@attribute MZ numeric
@attribute GX numeric
@attribute GY numeric
@attribute GZ numeric
@attribute TIME numeric
@attribute DATE numeric
@attribute SPEEDKMH numeric
@attribute LATITUDE numeric
@attribute LONGITUDE numeric
@attribute COMPORTAMENTO Pastando/Procurando,EmPe,Andando,Deitado

@data
A3,1157,9002,-12840,-460,1575,-205,243,195,-466,125307.770,290615,2026.5936,05443.3858,000.0,EmPe
A3,1226,9986,-12546,-463,1585,-264,-492,2209,-2189,125309.770,2026.5936,05443.3858,290615,000.0,EmPe
A3,870,8224,-13117,-527,1537,-248,590,-712,179,125318.770,290615,2026.5936,05443.3858,000.0,EmPe

```

5.8 Considerações Finais

Neste capítulo foram apresentados os detalhes da área, dos animais e da instalação dos colares. Foram apresentados os comportamentos observados, o procedimento utilizado para as anotações deles e os períodos de coleta, detalhando quais colares foram utilizados em cada animal e a quantidade de registros coletados.

As alterações propostas e realizadas no *software* Ápis foram apresentadas, além dos detalhes dos registros fornecidos por ele e o procedimento utilizado para unificá-los aos dados capturados nos colares, para posterior utilização na construção dos modelos de classificação do comportamento bovino.

6 Classificação dos Comportamentos

Neste capítulo são apresentados os resultados dos modelos propostos para aferição automática dos comportamentos. Na Seção 6.1 são apresentadas as ferramentas e os procedimentos adotados para a construção dos modelos de classificação. Na Seção 6.2 apresenta-se o procedimento utilizado para a classificação dos comportamentos primários (Pastando/Procurando, Andando, Em Pé, Deitado), a partir dos dados de movimentação e posicionamento do animal. Já a Seção 6.3 traz o procedimento de classificação dos comportamentos secundários, utilizando dados complementares, que classificaram os momentos que o animal estava Bebendo Água e Ruminando.

6.1 Processamento e Análise dos Dados

Nesta seção apresenta-se o *software* utilizado para a construção do modelo de classificação dos comportamentos, os algoritmos utilizados e o procedimento realizado para avaliar os conjuntos de dados utilizados na construção e validação do modelo proposto.

6.1.1 Ferramentas Utilizadas

O WEKA foi a ferramenta utilizada para construção e validação dos modelos de classificação. A versão utilizada foi a 3.6.12, que possui um pacote chamado *weka.classifiers*, com os algoritmos de classificação organizados nas categorias:

- *bayes*: classificadores Bayesianos;
- *functions*: classificadores Support Vector Machines, algoritmos de regressão, redes neurais;
- *lazy*: classificadores de vizinhos mais próximos;
- *meta*: classificadores que utilizam a base de um ou mais classificadores como entrada;
- *mi*: classificadores que lidam com dados de várias instâncias;
- *Misc*: classificadores que não se encaixam em qualquer outra categoria;
- *rules*: classificadores baseados em regras;
- *trees*: classificadores baseados em árvores de decisão.

Dos algoritmos de classificação do WEKA, 57 estavam aptos a processar os dados utilizados, considerando os padrões de entrada suportado por cada algoritmo. Alguns algoritmos, por exemplo, só suportam classes binárias, não sendo possível classificar clas-

ses multivariadas. A Tabela 14 apresenta os algoritmos utilizados para construção dos modelos.

Tabela 14 – Relação dos 57 algoritmos utilizados para construção dos modelos de classificação do comportamento animal.

Categoria	Algoritmo	Categoria	Algoritmo
bayes	BayesNet	meta	RandomCommittee
bayes	DMNBtext	meta	RandomSubSpace
bayes	NaiveBayes	meta	RotationForest
functions	Logistic	meta	Stacking
functions	MultilayerPerceptron	meta	StackingC
functions	RBFNetwork	meta	Vote
functions	SimpleLogistic	misc	HyperPipes
functions	SMO	misc	VFI
lazy	IB1	rules	ConjunctiveRule
lazy	IBk	rules	DecisionTable
lazy	KStar	rules	DTNB
lazy	LWL	rules	JRip
meta	AdaBoostM1	rules	NNge
meta	AttributeSelectedClassifier	rules	OneR
meta	Bagging	rules	PART
meta	ClassificationViaClustering	rules	Ridor
meta	ClassificationViaRegression	rules	ZeroR
meta	CVParameterSelection	trees	BFTree
meta	Dagging	trees	DecisionStump
meta	Decorate	trees	FT
meta	END	trees	J48
meta	FilteredClassifier	trees	J48graft
meta	Grading	trees	LADTree
meta	LogitBoost	trees	LMT
meta	MultiBoostAB	trees	RandomForest
meta	MultiClassClassifier	trees	RandomTree
meta	MultiScheme	trees	REPTree
meta	OrdinalClassClassifier	trees	SimpleCart
meta	RacedIncrementalLogitBoost		

6.1.2 Análise e Seleção dos Algoritmos

Para cada coleta realizada no campo foi gerado um arquivo ARFF, a partir da combinação dos registros dos colares e das observações. Os registros de todos os conjuntos foram agrupados em um único arquivo ARFF, que foi utilizado para avaliar cada um dos algoritmos. Para selecionar os algoritmos com melhor desempenho, foi avaliado a acurácia da classificação do modelo, utilizando o método *cross-validation* com 10 *folds*. Neste método o conjunto de dados é dividido em 10 partes de tamanhos aproximados, nos quais 9 partes são utilizadas como conjunto de treinamento e a parte restante como conjunto de teste. O processo repete-se 10 vezes, até que todos os *folds* sejam utilizados como conjunto de teste.

O arquivo ARFF, com todos os registros, possui 13.088 instâncias. Na Tabela 15, apresenta-se os detalhes individuais dos 10 conjuntos que formaram o arquivo geral. Cada conjunto representa um período de coleta realizado com um colar e um animal.

Na primeira fase dos testes, o modelo foi gerado para prever quatro comportamentos primários: Pastando/Procurando, Andando, Em Pé e Deitado. A Figura 36

Tabela 15 – Detalhes da quantidade de comportamentos de cada um dos conjuntos de dados utilizados para construção e validação do modelo de classificação.

Colar	Sequência	Animal	Comportamentos						Total
			Andando	Pastando/Procurando	Deitado		Em Pé		
					Ruminando	Parado	Ruminando	Parado	
A	2	6998	88	745	174	37	3	65	1112
A	3	6998	271	453	468	311	285	245	2033
B	2	7014	83	779	118	23	24	104	1131
B	3	7135	124	675	117	438	131	250	1735
C	3	6998	127	524	545	276	184	196	1852
C	4	6998	38	229	85	6	0	48	406
D	1	7014	154	545	75	127	51	174	1126
D	2	7014	164	914	103	137	141	231	1690
D	3	7135	138	652	316	262	13	217	1598
D	4	7135	42	216	33	85	0	29	405
		Total	1229	5732	2034	1702	832	1559	13088

apresenta o esquema para identificação dos comportamentos primários, em que são utilizados os dados dos sensores de movimentação (acelerômetro, giroscópio e magnetômetro) e os dados de posicionamento (GPS). Os comportamentos "Em Pé-Parado" e "Em Pé-Ruminando", foram agrupados em uma única classe, definida como "Em Pé"; o mesmo foi realizado com as classes "Deitado-Parado" e "Deitado-Ruminando", que foram agrupadas na classe Deitado. Na Seção 6.3.1 são apresentados os resultados da classificação incluindo os comportamentos "Em Pé-Ruminando" e "Deitado-Ruminando".

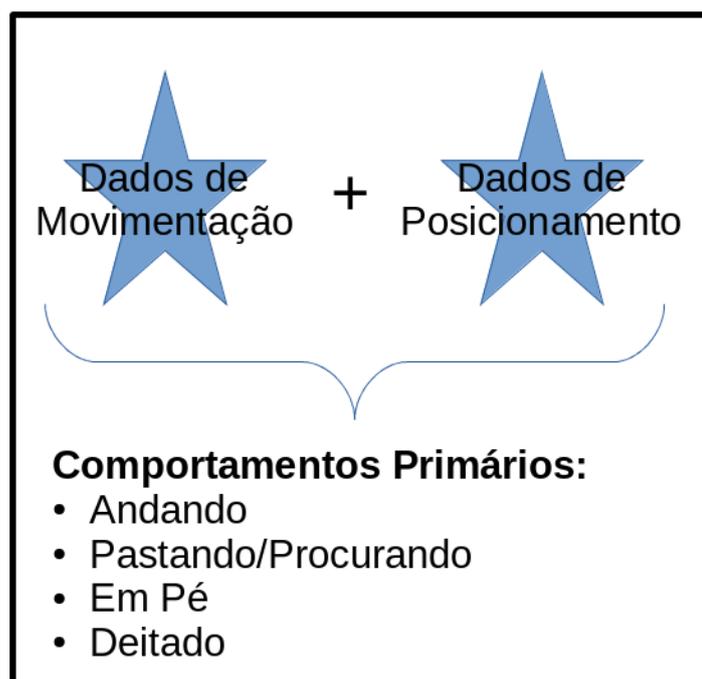


Figura 36 – Comportamentos primários identificados a partir dos dados de posição e movimentação.

Para automatizar a tarefa de construção e validação dos modelos nos 57 algoritmos citados, foi criado um código Java que gerou cada um dos modelos e realizou a validação. Um arquivo de saída gravou o resultado, apresentando o algoritmo utilizado e a acurácia obtida na validação do modelo. Todos eles foram executados utilizando os parâmetros padrões do WEKA.

Dentre os algoritmos avaliados, 19 apresentaram acurácia acima de 90% e 10 deles

acurácia abaixo de 50%. Os algoritmos *Random Forest*, *Rotation Forest*, *Random Committee*, *END* e *Bagging* apresentaram os melhores valores. Na Tabela 16, são apresentados os 10 algoritmos com melhores resultados. O conjunto de dados utilizado agrupou todos os registros disponíveis.

Tabela 16 – Algoritmos com melhor acurácia na validação utilizando o método *cross-validation* com 10 *folds* com todos os registros.

Categoria	Algoritmo	Acurácia (%)
trees	Random Forest	93,9868
meta	Rotation Forest	93,7576
meta	Random Committee	93,4749
meta	END	93,0012
meta	Bagging	92,7338
meta	Decorate	92,4587
lazy	IB1	91,8627
lazy	IBk	91,8627
trees	J48 graft	91,4349
meta	Classification Via Regression	91,4043

6.1.3 Avaliação dos Conjuntos de Dados Individuais

Para avaliar cada conjunto de dados individual (cada conjunto representa um período de coleta obtida com um colar e um animal específico), foi criado um arquivo ARFF de teste, composto por cerca de 15% dos registros de cada conjunto, totalizando 1.975 instâncias. Os registros foram selecionados aleatoriamente nos arquivos, com o cuidado apenas para que as classes não ficassem desbalanceadas. A Tabela 17 apresenta o número de registros de cada classe no arquivo de teste.

Tabela 17 – Detalhes do arquivo de teste com 15% dos dados de cada conjunto.

Classe	Nº de Instâncias
Andando	201
Pastando/Procurando	851
Em Pé	359
Deitado	564
Total	1975

Para validar cada conjunto foi utilizado o método *Holdout* (apresentado na Seção 2.4.6), utilizando o arquivo do conjunto para criar o modelo e os registros do arquivo teste para validação. Desta maneira todos os modelos foram validados com o mesmo conjunto de teste.

Para os conjuntos de dados utilizados para construir os modelos (apresentados na sequência), foram descartados os 15% dos dados utilizados no arquivo de teste.

Os 5 algoritmos com melhor resultado na validação anterior (*Random Forest*, *Rotation Forest*, *Random Committee*, *END* e *Bagging*) foram utilizados para estes testes. A Tabela 18 apresenta a acurácia da validação de cada conjunto com os cinco algoritmos.

Tabela 18 – Resultados da validação individual dos conjuntos utilizando o método *Holdout*.

Conjunto	Random Forest	Rotation Forest	Random Committee	END	Bagging
A2	78,33	76,66	74,38	79,24	82,68
A3	78,99	78,23	79,54	76,76	75,44
B2	67,85	66,99	64,20	65,52	66,33
B3	82,18	80,91	81,11	81,77	80,71
C3	78,28	78,63	76,20	78,33	77,32
C4	50,33	53,16	51,24	53,92	56,46
D1	84,35	83,85	79,70	80,86	82,68
D2	83,14	81,22	82,53	79,04	81,22
D3	78,48	78,53	75,80	75,59	76,00
D4	73,37	69,32	59,24	74,38	79,49

Os conjuntos B2 e C4 apresentaram as menores acurácias. Para avaliar em detalhes os resultados dos modelos construídos, analisou-se a matriz de confusão de cada um deles, gerada com o algoritmo *Random Forest*.

Avaliando o resultado do modelo criado com o conjunto B2, observa-se que a classe Deitado possui a menor acurácia (8,7%), sendo que 81,2% dos registros desta classe foram classificados como Em Pé, como pode ser visto na Tabela 19. As classes Em Pé e Deitado são aquelas em que a posição do pescoço do animal são mais parecidos, registrando pouca variação na posição. Observando os registros que formam o conjunto B2 (Tabela 15), identifica-se que há poucos registros, que não são suficientes para construir um modelo que classifique os quatro comportamentos com valores aceitáveis.

Tabela 19 – Matriz de confusão do modelo gerado pelo conjunto B2 e o algoritmo *Random Forest* e validado com o método *Holdout*.

=== Matriz de Confusão ===					
a	b	c	d	<- classificado como	Acurácia (%)
819	14	18	0	a = Pastando/Procurando	96,2
66	289	3	1	b = Em Pé	80,5
16	2	183	0	c = Andando	91
56	458	1	49	d = Deitado	8,7

Já no resultado do modelo do conjunto C4, apenas a classe Pastando/Procurando possui resultados aceitáveis. Avaliando o conjunto de dados identificou-se que há poucos registros rotulados como Andando, Em Pé e Deitado, fazendo com que a classificação destes comportamentos tenha baixa acurácia. A matriz de confusão gerada na classificação do modelo do conjunto C4 é apresentado na Tabela 20.

Tabela 20 – Matriz de confusão do modelo gerado pelo conjunto C4 e o algoritmo *Random Forest* e validado com o método *Holdout*.

=== Matriz de Confusão ===					
a	b	c	d	<- classificado como	Acurácia (%)
847	0	4	0	a = Pastando/Procurando	99,5
297	61	1	0	b = Em Pé	17
142	1	58	0	c = Andando	28,9
173	363	0	28	d = Deitado	5

6.1.4 Avaliação de um Conjunto Descartado

O conjunto A1 foi descartado na construção do modelo pelo fato do colar ter sido colocado invertido no animal (conforme relatado na Seção 5.6). Como foram utilizados os valores absolutos registrados pelo sensores, a posição do colar interfere na posição de cada um dos três eixos dos sensores de movimentação. A validação se deu no intuito de confirmar os problemas relatados quando o equipamento é colocado incorretamente no animal.

A acurácia média obtida com os cinco algoritmos (*Random Forest*, *Rotation Forest*, *Random Committee*, *END* e *Bagging*) foi de apenas 19,74%. Analisando as classes individualmente, no modelo gerado com o algoritmo *Random Forest*, nota-se que a classe Pastando/Procurando foi a que obteve melhor resultado, com pouco mais de 36% dos registros classificados corretamente, e a classe Deitado o pior resultado, menos de 1% de acurácia. A Tabela 21 apresenta a matriz de confusão da classificação.

Tabela 21 – Matriz de confusão com percentual de acerto da validação do modelo gerado com o conjunto A1 e o algoritmo *Random Forest*.

=== Matriz de Confusão ===					
a	b	c	d	<- classificado como	Acurácia (%)
313	329	28	181	a = Pastando/Procurando	36,8
308	4	0	47	b = Em Pé	1,1
156	0	41	4	c = Andando	20,4
528	31	0	5	d = Deitado	0,9

Comparado aos resultados dos demais conjuntos, confirma-se que a posição do equipamento no animal interfere diretamente no resultado da classificação.

A posição do colar não interfere nas coletas do GPS. Como o modelo aqui proposto utiliza apenas parte dos dados do GPS, não é possível somente com eles aferir as classes desejadas. Os trabalhos que aferiram os comportamentos somente com o GPS, utilizaram variáveis não utilizadas neste trabalho e geraram outros atributos a partir da combinação dos dados coletados, como pode-se observar no trabalho de Jesus (2014).

6.1.5 Combinação dos Conjuntos de Dados

Para avaliar o impacto/variação que cada colar e animal possui no modelo construído, agrupou-se os dados em novos conjuntos, chamados de conjuntos combinados. Eles foram gerados, a partir dos conjuntos individuais, para validar os algoritmos com dados de diferentes situações, que foram baseados em dois critérios:

1. Conjuntos de dados agrupados por animal: três animais foram utilizados nos experimentos em campo;
2. Conjuntos de dados agrupados por colar: quatro colares foram utilizados nas coletas em campo;

As combinações foram feitas de maneira que todas as possibilidades de variação dos conjuntos fossem atendidas, apresentando a acurácia obtida com o modelo gerado em cada cada combinação. A partir da comparação dos resultados pode-se identificar o impacto que os dados de cada colar ou animal possuem no modelo gerado.

Os conjuntos de dados utilizados no agrupamento foram aqueles já com os 15% dos dados retirados para o conjunto de teste. Este conjunto de teste foi novamente utilizado para validar os modelos, que foram criados com o algoritmo *Random Forest*.

Os resultados da validação, utilizando as combinações dos animais e dos colares, são apresentados a seguir.

6.1.5.1 Variação dos Dados por Animal

Os conjuntos individuais foram agrupados aos respectivos animais utilizados nas coletas. A Tabela 22 apresenta os conjuntos que foram agrupados por animal e o número total de instâncias em cada um.

Tabela 22 – Conjuntos agrupados por animal a partir dos dados dos conjuntos individuais.

Animal	Conjuntos	Nº de Instâncias
6998	A2, A3, C3, C4	4588
7014	B2, D1, D2	3351
7135	B3, D3, D4	3174

Os conjuntos dos animais foram combinados de maneira que sete modelos de classificação fossem gerados. A seguir são apresentadas as combinações geradas com os dados para construir os modelos:

- Combinação 1: Animal 6998
- Combinação 2: Animal 7014
- Combinação 3: Animal 7135
- Combinação 4: Animal 6998 + Animal 7014
- Combinação 5: Animal 6998 + Animal 7135
- Combinação 6: Animal 7014 + Animal 7135
- Combinação 7: Animal 6998 + Animal 7014 + Animal 7135

O resultado da validação de cada modelo é apresentado na Tabela 23. O conjunto que agrupou os dados dos três animais obteve o melhor resultado de classificação, indicando que quanto maior o conjunto de dados maior os detalhes sobre as classes aferidas.

6.1.5.2 Variação dos Dados por Colar

Desta vez os conjuntos foram agrupados aos respectivos colares utilizados nas coletas. A Tabela 24 apresenta os conjuntos agrupados e o número total de instâncias em

Tabela 23 – Resultado da validação dos conjuntos agrupados por animal utilizando o método *Holdout*.

Combinação	Acurácia (%)
Animal 6998+7014+7135	94,7341
Animal 6998+7014	92
Animal 6998+7135	90,9367
Animal 6998	87,8481
Animal 7014+7135	87,3417
Animal 7014	83,3417
Animal 7135	82,4810

cada um.

Tabela 24 – Conjuntos agrupados por colar a partir dos dados dos conjuntos individuais.

Colar	Conjuntos	Nº de Instâncias
A	A2, A3	2671
B	B2, B3	2434
C	C3, C4	1917
D	D1, D2, D3, D4	4091

Os conjuntos foram combinados de maneira que quinze modelos fossem gerados. A seguir são apresentadas as combinações utilizadas:

- Combinação 1: Colar A
- Combinação 2: Colar B
- Combinação 3: Colar C
- Combinação 4: Colar D
- Combinação 5: Colar A + Colar B
- Combinação 6: Colar A + Colar C
- Combinação 7: Colar A + Colar D
- Combinação 8: Colar B + Colar C
- Combinação 9: Colar B + Colar D
- Combinação 10: Colar C + Colar D
- Combinação 11: Colar A + Colar B + Colar C
- Combinação 12: Colar A + Colar B + Colar D
- Combinação 13: Colar A + Colar C + Colar D
- Combinação 14: Colar B + Colar C + Colar D
- Combinação 15: Colar A + Colar B + Colar C + Colar D

A Tabela 25 apresenta a acurácia obtida com a validação de cada modelo. O melhor resultado foi obtido quando utilizou-se os dados de todos os colares, resultado similar ao apresentado na variação dos dados por animal.

Tabela 25 – Resultado da validação dos conjuntos agrupados por colar utilizando o método *Holdout*.

Combinação	Acurácia (%)
Colar A+B+C+D	94,7341
Colar A+B+D	92,7088
Colar A+C+D	92,7088
Colar A+B+C	91,2405
Colar A+D	90,6329
Colar B+C+D	88,2025
Colar A+C	87,8481
Colar B+D	87,3417
Colar A+B	87,2911
Colar D	86,4810
Colar C+D	85,9746
Colar B+C	83,4430
Colar B	81,9240
Colar A	79,3924
Colar C	76,3544

6.1.5.3 Análise dos Resultados das Variações

Observa-se que conforme foram agregados os conjuntos dos animais, a acurácia do modelo foi aumentando. A exceção foi o conjunto do animal 6998 que apresentou acurácia melhor que os outros dois conjuntos, inclusive na ocasião em que eles foram agregados (Animal 7014 + 7135).

O impacto dos conjuntos por colar aumentou gradativamente conforme os conjuntos foram agregados. Com exceção do conjunto D, que apresentou melhor resultado que os demais colares (Colar A, Colar B e Colar C), inclusive que os conjuntos agregados entre Colar B+C e Colar C+D.

Observou-se que os conjuntos individuais com melhores resultados são os que possuem mais registros (Colar D e Animal 6998) e que o conjunto com todos os registros é o que possui a melhor acurácia.

A variação mostra que quando os conjuntos de novos animais ou colares são adicionados o resultado é melhor, pois maior é o conjunto de treinamento utilizado. Na validação dos conjuntos por animal, observa-se uma variação de mais de 2% quando os dados de um dos animais é retirado. Vale ressaltar que o conjunto de teste está limitado a três animais, sendo que para um conjunto com mais animais o impacto pode ser diferente.

Para a construção e validação dos modelos foi utilizado o conjunto com todos os registros, e os resultados são apresentados nas seções a seguir.

6.2 Modelos de Classificação dos Comportamentos Primários

O modelo de classificação dos comportamentos primários foi construído com o conjunto de dados que agregou os registros de todos os conjuntos. Inicialmente apresentava-se o resultado da validação do modelo construído com o algoritmo *Random Forest*. Na sequência os resultados do classificador construído com o algoritmo SVM, que não estava

na lista inicial dos algoritmos utilizados.

6.2.1 Validação do Modelo Utilizando o Algoritmo Random Forest

A validação do modelo, utilizando o método *cross-validation* com 10 *folds*, apresentou acurácia média de 93,99%. Em um universo de 13.088 instâncias, 12.301 foram classificadas corretamente e 787 incorretamente. Observa-se que a classe com a menor acurácia é a do comportamento Em Pé, com 88,6%. A classe Deitado apresenta 93,4%, Andando 95% e Pastando/Procurando 96,4%.

Na Tabela 26 é apresentada a matriz de confusão do modelo, mostrando detalhes do número de instâncias classificadas correta e incorretamente em cada classe. Analisando a classe Em Pé, que possui menor acurácia, observa-se que 95,9% dos registros incorretamente classificados estão nas classes Deitado e Pastando/Procurando.

Tabela 26 – Matriz de confusão da validação do modelo de classificação dos comportamentos primários utilizando o método *cross-validation*.

=== Matriz de Confusão ===					
a	b	c	d	<- classificado como	Acurácia (%)
5524	87	105	16	a = Pastando/Procurando	96,4
124	2119	11	137	b = Em Pé	88,6
52	7	1168	2	c = Andando	95
67	177	2	3490	d = Deitado	93,4

6.2.2 Validação do Modelo Utilizando o Algoritmo SVM

Apesar do SVM não estar na lista inicial de algoritmos apresentados, por exigir a configuração de alguns pacotes no WEKA, ele é comumente citado na literatura e utilizado em problemas de classificação. Assim, configurou-se o WEKA para que o algoritmo fosse executado e o modelo por ele construído avaliado.

Novamente o conjunto com todos os dados foi utilizado junto com o método de validação *cross-validation* com 10 *folds*. Como nos outros algoritmos, inicialmente utilizou-se a configuração padrão do WEKA. O resultado inicial foi de apenas 43,79% de acurácia.

Considerando as possibilidades de ajustes nas configurações do *kernel* utilizado no SVM, os parâmetros C (*cost*) e γ (*gamma*) foram ajustados, utilizando o método *grid search*, que busca pelos melhores parâmetros testando sequências exponenciais e analisando os resultados obtidos com a execução do algoritmo (HSU; CHANG; LIN, 2010). O *kernel* utilizado foi o *radial basis function* (RBF). Com o ajuste dos parâmetros, o classificador apresentou acurácia de 93,2%. Do total de 13.088 registros, 12.198 foram corretamente classificados, enquanto 890 foram classificadas incorretamente .

A Tabela 27 compara o resultado obtido ao apresentado com o modelo do algoritmo *Random Forest*. A acurácia média apresenta diferença de menos de 1%. Avaliando individualmente as classes nota-se que os comportamentos Em Pé e Deitado tiveram melhor

desempenho no SVM, com variação de 1,5% e 2,1% respectivamente. Já os comportamentos Pastando/Procurando e Andando tiveram desempenho inferior, com redução de -2,1% e -7,9% respectivamente.

Tabela 27 – Comparação da acurácia de cada classe nos modelos gerados com os algoritmos *Random Forest* e SVM.

Classe	Random Forest	SVM	Diferença
Pastando/Procurando	96,4%	94,3%	2,1%
Em Pé	88,6%	90,1%	-1,5%
Andando	95%	87,1%	7,9%
Deitado	93,4%	95,5%	-2,1%

6.3 Modelos de Classificação dos Comportamentos Secundários

Com o modelo para classificação dos quatro comportamentos primários validado, buscou-se a identificação automática de novos comportamentos. A proposta foi identificar comportamentos com auxílio de variáveis complementares.

Estes novos comportamentos são aqui definidos como comportamentos secundários e caracterizados como aqueles identificados com o uso dos dados do GPS, dos sensores de movimentação e alguma outra variável. A Figura 37 mostra o esquema da identificação dos comportamentos secundários, apresentando algumas das possíveis variáveis: locais de interesse (georreferenciados), dados do ambiente (sensor de luminosidade, temperatura, etc) e as próprias classes dos comportamentos primários. Por exemplo, para identificar se o animal está bebendo água, além dos dados de movimentação indicarem que ele está com a cabeça baixa, pode-se utilizar a posição georreferenciada do bebedouro, para confirmar se o animal estava naquele local; para o comportamento Ruminando, pode-se verificar antes se ele está Em Pé ou Deitado, comportamentos primários que se caracterizam quando o animal está ruminando.

Os comportamentos secundários abordados foram: Ruminando e Bebendo Água. Os resultados da classificação dos modelos, para predição destes comportamentos, são apresentados a seguir.

6.3.1 Comportamento Ruminando

O comportamento Ruminando se caracteriza pelo movimento da mandíbula do animal (movimento de mastigação) e ocorre quando ele está Parado Em Pé ou Deitado, ou seja, se o animal está em um destes dois comportamentos, pode-se verificar se ele está ruminando ou não. Neste modelo utilizou-se as classes dos comportamentos primários e os dados dos sensores de movimentação e de posicionamento do animal, conforme esquema apresentado na Figura 38. Para padronizar os termos serão utilizadas as classes "Em Pé-



Figura 37 – Proposta de modelo para classificação de comportamentos secundários.

Ruminando" e "Deitado-Ruminando" para os momentos que o animal está ruminando e "Em Pé-Parado" e "Deitado-Parado" para quando o animal está parado.

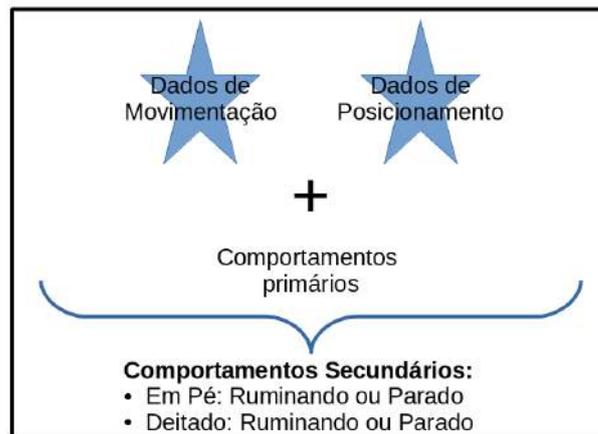


Figura 38 – Esquema para classificação do comportamento Ruminando.

Dois testes foram realizados para a identificação dos comportamentos: no primeiro os comportamentos foram classificados juntamente com os comportamentos primários, ou seja, o modelo predizia seis possíveis comportamentos (Pastando/Procurando, Andando, Em Pé-Ruminando, Deitado-Ruminando, Em Pé-Parado e Deitado-Parado), sem distinguir os comportamentos primários e secundários; no segundo teste apenas os registros classificados como Em Pé e Deitado foram utilizados para indicar uma segunda classificação, mostrando se o animal estava ou não ruminando.

O algoritmo *Random Forest* foi utilizado em ambos os testes, com o método de validação *cross-validation* com 10 *folds* e com o conjunto de dados formado por todos os registros coletados. O conjunto, composto por 13088 instâncias, possui 2034 registros do animal Deitado-Ruminando e 832 registros como Em Pé-Ruminando.

6.3.1.1 Classificação do Comportamento Ruminando como Comportamento Primário

A acurácia média do modelo foi de 89,55%, considerado um bom resultado para um modelo de classificação. Porém, analisando individualmente as classes, constata-se baixa acurácia na classe Em Pé-Parado e Em Pé-Ruminando.

O modelo obteve acurácia de 79,2% na classe Em Pé-Parado e 67,9% na classe Em Pé-Ruminando. As classes Deitado-Parado e Deitado-Ruminando obtiveram acurácia de 90,8% e 80,4%, respectivamente. A Tabela 28 apresenta os detalhes da matriz de confusão do modelo de classificação.

Tabela 28 – Matriz de confusão do modelo de classificação do comportamento ruminando junto aos comportamentos primários.

=== Matriz de Confusão ===							
a	b	c	d	e	f	<- classificado como	Acurácia (%)
5540	57	9	107	9	10	a = Pastando/Procurando	96,7
119	1234	107	5	68	26	b = Em Pé-Parado	79,2
33	196	565	6	30	2	c = Em Pé-Ruminando	67,9
55	4	1	1167	1	1	d = Andando	95
15	46	31	1	1846	95	e = Deitado-Ruminando	90,8
53	82	13	1	184	1369	f = Deitado-Parado	80,4

Avaliando o número de instâncias corretamente classificadas nas classes Em Pé-Parado e Em Pé-Ruminando e comparando os resultados com a validação do modelo com os quatro comportamentos primários (em que se agrupou os dois comportamentos na classe Em Pé), observa-se uma redução da taxa de acerto. No modelo com quatro comportamentos, a acurácia da classe Em Pé era de 88,6% e com o comportamento Ruminando foi de 73,55%.

No mesmo comparativo, agora com a classe Deitado, os resultados mostram que a média foi de 85,6%, com o comportamento Ruminando, enquanto com o modelo com quatro comportamentos primários era de 93,4%.

6.3.1.2 Classificação do Comportamento Em Pé-Ruminando

Utilizando a abordagem da classificação como comportamento secundário, utilizou-se apenas os registros rotulados como Em Pé, ignorando aqueles marcados como Andando, Pastando/Procurando e Deitado.

O comportamento Em Pé foi então classificado em duas classes: Parado e Ruminando. O conjunto de dados com estas classes possui 2.391 registros, sendo 832 rotulados como Em Pé-Ruminando e 1.559 como Em Pé-Parado.

A classe Em Pé-Ruminando obteve 71,3% de acurácia e a classe Em Pé-Parado 91,1%, uma média de 81,2%. A Tabela 29 mostra a matriz de confusão do modelo.

No resultado da classificação, precisa-se considerar a taxa de erro da classificação dos comportamentos primários (primeiro classifica-se os comportamentos primários e de-

Tabela 29 – Matriz de confusão do modelo de classificação do comportamento Em Pé-Ruminando.

=== Matriz de Confusão ===			
a	b	<- classificado como	Acurácia (%)
1421	138	a = Em Pé-Parado	91,1
239	593	b = Em Pé-Ruminando	71,3

pois o comportamento ruminando), em que há uma taxa de erro de 11,4% no classe Em Pé. Considerando-a, a acurácia do modelo passa para 71,94%, sendo 63,17% na classe Em Pé-Ruminando e 80,71% na classe Em Pé-Parado.

6.3.1.3 Classificação do Comportamento Deitado-Ruminando

Utilizou-se a mesma abordagem da validação do comportamento Em Pé-Ruminando, descartando os registros rotulados como Andando, Pastando/Procurando e Em Pé.

O comportamento Deitado também foi classificado em duas classes: Parado e Ruminando. O conjunto de dados com estas classes possui 3.736 registros, sendo 2.034 rotulados como Deitado-Ruminando e 1.702 como Deitado-Parado.

A classe Deitado-Ruminando obteve 92% de acurácia e a classe Deitado-Parado 85,2%, uma média de 88,6%. A Tabela 30 mostra a matriz de confusão do modelo.

Tabela 30 – Matriz de confusão do modelo de classificação do comportamento Deitado-Ruminando.

=== Matriz de Confusão ===			
a	b	<- classificado como	Acurácia (%)
1871	163	a = Deitado-Ruminando	92
252	1450	b = Deitado-Parado	85,2

Considerando a taxa de erro da classificação dos comportamentos primários, a classe Deitado possui 6,6% de erro, ficando o percentual de acerto da classe Deitado-Ruminando em 85,92% e da classe Deitado-Parado em 79,57%, uma média de 82,75%.

6.3.1.4 Resultados da Classificação do Comportamento Ruminando

As duas abordagens utilizadas para classificar o comportamento Ruminando - classificando junto aos comportamentos primários e classificando em duas etapas - apresentaram resultados parecidos.

A Tabela 31 apresenta o resumo da classificação dos comportamentos Em Pé-Ruminando e Em Pé-Parado, comparando a acurácia do modelo de classificação que inclui o comportamento ruminando como comportamento primário e o modelo de classificação em duas etapas (primeiro classifica os comportamentos primários e depois classifica o comportamento ruminando).

Os resultados das classificações dos comportamentos Deitado-Ruminando e Deitado-Parado, nos dois modelos propostos, são apresentados na Tabela 32.

Tabela 31 – Acurácia da classificação dos comportamentos Em Pé-Ruminando e Em Pé-Parado.

	Ruminando como Primário	Duas etapas (Primários + Secundários)
Em Pé-Parado	79,2%	80,71%
Em Pé-Ruminando	67,9%	63,17%
Média	73,55%	71,94%

Tabela 32 – Acurácia da classificação dos comportamentos Deitado-Ruminando e Deitado-Parado.

	Ruminando como Primário	Duas etapas (Primários + Secundários)
Deitado-Parado	80,4	79,57
Deitado-Ruminando	90,8	85,92
Média	85,6	82,75

O modelo que classificou o comportamento Ruminando junto aos comportamentos primários apresentou melhor acurácia tanto para os momentos que o animal estava Em Pé quanto Deitado.

Deve-se considerar nestes resultados os erros das observações em campo, pois um dos problemas identificados, na observação dos comportamentos, foi diferenciar os momentos em que o animal estava parado ou ruminando, pois dependendo da posição que ele estava, não era possível identificar se havia movimento de mastigação, característica que identifica o comportamento ruminando.

Outra observação é quanto ao movimento que caracteriza o comportamento ruminando. A variação da posição do pescoço do animal, comparando os comportamentos parado e ruminando, é pequena, podendo dificultar a distinção dos dois comportamentos.

6.3.2 Comportamento Bebendo Água

Para aferir o comportamento Bebendo Água, incluiu-se os registros que foram rotulados com este comportamento nas observações. Há 513 instâncias rotuladas como Bebendo Água, que foram agregadas ao conjunto de dados com todos os registros.

Duas estratégias foram utilizadas para a construção do modelo de classificação. Na primeira, o comportamento Bebendo Água é classificado junto aos comportamentos primários, e na segunda, classificado como um comportamento secundário. Para construção de todos os modelos apresentados a seguir, foi utilizado o algoritmo *Random Forest* e validado com o método *cross-validation* com 10 *folds*.

6.3.2.1 Classificação do Comportamento Bebendo Água como Comportamento Primário

Inicialmente, para a classificação do comportamento junto aos comportamentos primários, foram utilizados os dados dos sensores de movimentação e a velocidade de deslocamento registrada pelo GPS. O comportamento Bebendo Água obteve 71,2% de acurácia.

O movimento e posição do animal no comportamento Bebendo Água é caracterizado pelo cabeça para baixo, similar as características do comportamento Pastando/Procurando. Observa-se na matriz de confusão do modelo, apresentada na Tabela 33, que a maior taxa de erro da classificação do comportamento é com a classe Pastando/Procurando (23,39% dos registros).

Tabela 33 – Matriz de confusão com acurácia do modelo de classificação do comportamento Bebendo Água como comportamento primário.

==== Matriz de Confusão ====						
a	b	c	d	e	classificado como	Acurácia (%)
5529	83	98	18	4	a = Pastando/Procurando	96,5
125	2119	10	130	7	b = Em Pé	88,6
58	5	1164	2	0	c = Andando	94,7
60	166	3	3503	4	d = Deitado	93,8
120	21	4	3	365	e = Bebendo Água	71,2

Uma característica que diferencia o comportamento Bebendo Água dos demais é a localização do animal no momento em que o comportamento acontece, considerando a posição fixa dos bebedouros na área de pasto. Para os outros comportamentos investigados neste trabalho não se pode determinar que sempre ocorrem em um determinado local.

O GPS coleta as coordenadas geográficas (latitude e longitude) do animal, que podem ser utilizadas na classificação dos momentos que ele estava bebendo água. Assim, a latitude e longitude foram incluídas aos registros e um novo modelo foi criado para classificação. Nele o comportamento Bebendo Água obteve 89,1% de acurácia, conforme apresentado na Tabela 34.

Tabela 34 – Matriz de confusão da classificação dos comportamentos primários e Bebendo Água, utilizando as coordenadas geográficas.

==== Matriz de Confusão ====						
a	b	c	d	e	classificado como	Acurácia (%)
5553	68	94	16	1	a = Pastando/Procurando	96,9
107	2229	11	38	6	b = Em Pé	93,2
49	5	1173	2	0	c = Andando	95,4
53	61	2	3620	0	d = Deitado	96,9
34	17	5	0	457	e = Bebendo Água	89,1

A utilização das coordenadas geográficas na classificação apresenta um problema: o modelo fica restrito a dados coletados na mesma área dos dados utilizados na construção do modelo, ou seja, ele não pode ser generalizado para dados coletados em outras áreas. O modelo proposto para classificação dos Comportamentos Primários na Seção 6.2 não possui esta restrição, pois a latitude e longitude não foram utilizadas.

Já na classificação do comportamento Bebendo Água, as coordenadas são fundamentais, já que a localização que o animal está (área do bebedouro) é um dos principais atributos para identificar o comportamento. A variação entre o modelo que utilizou (89,1%) e não utilizou (71,2%) as coordenadas é de 17,9%.

6.3.2.2 Classificação do Comportamento Bebendo Água como Comportamento Secundário

Como o interesse neste modelo é o de identificar os momentos que o animal está bebendo água, agrupou-se os registros em duas classes: Bebendo Água e Outros. A classe Outros foi formada por todos os registros que não estavam rotulados como Bebendo Água. Utilizou-se os dados dos sensores de movimentação, as coordenadas geográficas e a velocidade de deslocamento.

A Tabela 35 apresenta a matriz de confusão do modelo. O resultado mostra um aumento de 3,9% na acurácia da classificação da classe Bebendo Água, se comparado ao resultado da classificação como comportamento primário.

Tabela 35 – Matriz de confusão da classificação do comportamento Bebendo Água agregando os demais comportamentos em uma única classe.

=== Matriz de Confusão ===			
a	b	classificado como	Acurácia (%)
13077	11	a = Outro	99,9
36	477	b = Bebendo Água	93

O esquema da Figura 39 apresenta a proposta de classificação dos comportamentos utilizando áreas de interesse, além dos dados coletados pelos sensores.

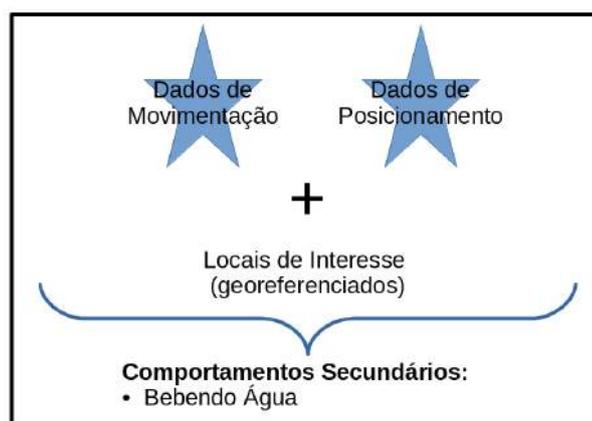


Figura 39 – Esquema para classificação do comportamento Bebendo Água.

Assim, um novo modelo foi construído, agora utilizando apenas os registros que foram gravados na área próxima ao bebedouro. Para limitar a área, analisou-se na ferramenta WEKA o gráfico que plota os atributos e selecionou-se as coordenadas que estavam nos limites dos registros gravados como Bebendo Água. Na Figura 40 são apresentados os pontos dos registros rotulados como Bebendo Água (pontos em negrito e em destaque). A área possui dois bebedouros, por isso é possível visualizar duas áreas em destaque. Os pontos marcados em cor mais clara indicam a localização do animal quando identificado em outros comportamentos.

O conjunto com os registros da área selecionada possui 1.224 registros, sendo 711 classificados como Outros e 513 como Bebendo Água. O modelo criado com eles apresentou

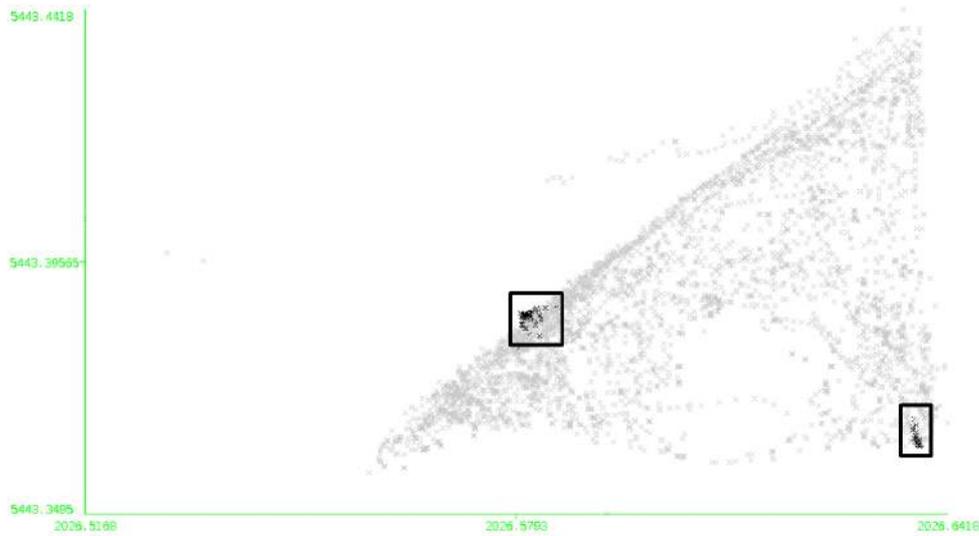


Figura 40 – Plotagem da localização dos animais na área de pasto baseado nas coordenadas geográficas: em destaque os registros rotulados como Bebendo Água.

acurácia de 94,1% no comportamento Bebendo Água, que representa um aumento de 1,1% na comparação com o modelo que utiliza todos os registros. A Tabela 36 apresenta os resultados da nova classificação, utilizando os dados filtrados.

Tabela 36 – Matriz de confusão da classificação do comportamento Bebendo Água apenas com os registros filtrados na área do bebedouro.

=== Matriz de Confusão ===			
a	b	classificado como	Acurácia (%)
701	10	a = Outro	98,6
30	483	b = BebendoÁgua	94,1

6.3.2.3 Resultados da Classificação do Comportamento Bebendo Água

Para a classificação do comportamento Bebendo Água foram utilizadas quatro estratégias diferentes:

1. Classificação da classe como Comportamento Primário;
2. Classificação da classe como Comportamento Primário acrescentando os dados de latitude e longitude ao conjunto de dados utilizado;
3. Agrupando os comportamentos sem interesse em uma única classe;
4. Agrupando os comportamentos sem interesse em uma única classe e utilizando apenas os registros que estavam próximos a área de interesse.

A Tabela 37 apresenta um resumo da acurácia obtida na classificação do comportamento em cada uma das estratégias. Observa-se que a maior variação se dá quando os dados da coordenadas geográficas são utilizadas, enquanto as técnicas que agruparam os registros em apenas duas classes e filtraram os registros, apresentaram melhoras menos significativas na acurácia da classe de interesse.

Tabela 37 – Comparativo da acurácia nas quatro estratégias utilizadas para classificação do comportamento Bebendo Água.

Comportamento Primário	Comportamento Primário com Coord.Geográficas	Comportamentos Agrupados na Classe Outros	Comportamentos Agrupados Filtrados na Área de Interesse
71,2%	89,1%	93%	94,1%

Vale ressaltar que a utilização das coordenadas geográficas fazem com que o modelo fique restrito à classificação utilizando dados coletados em uma única área (mesma área dos dados utilizados na construção do modelo). Porém, mostrou-se eficiente na classificação do comportamento que possui relação direta com o posicionamento do animal na área do pasto, como o caso do comportamento Bebendo Água.

6.4 Comparação dos Resultados com Outros Trabalhos

Os resultados dos comparativos apresentados servem apenas como base para indicar os valores obtidos neste trabalho e os apresentados por outros. Considerando que os conjuntos de testes são diferentes, assim como o número de animais e o tempo de observação, não é justo afirmar que os resultados de um ou outro são melhores.

O trabalho de [Martiskainen et al. \(2009\)](#) classificou comportamentos, utilizando o algoritmo SVM e um conjunto de dados com 11.693 registros obtidos de colares com acelerômetro triaxial de 30 vacas. Os resultados apresentados no trabalho são de uma bateria de testes realizado com 1.000 instâncias. Para efeito de comparativo, as classes Em Pé-Ruminando e Deitado-Ruminando foram agrupadas na classe Ruminando e um novo modelo foi construído e validado.

A Tabela 38 compara a acurácia de cada classe entre o trabalho de [Martiskainen et al. \(2009\)](#) e os obtidos com este novo modelo. Destaque para a classe Deitado, que foi a única no trabalho de [Martiskainen et al. \(2009\)](#), que obteve melhor acurácia que o modelo aqui proposto.

Tabela 38 – Comparativo dos resultados obtidos com os apresentados por [Martiskainen et al. \(2009\)](#).

Classe	Acurácia	
	Martiskainen et al. (2009)	Esta pesquisa
Pastando	81%	96,5%
Andando	79%	94,9%
Em Pé	65%	75,8%
Deitado	83%	78,7%
Ruminando	86%	87,9%

[Nadimi et al. \(2012\)](#) também utilizou um acelerômetro preso em um colar, mas avaliou o comportamento de ovelhas. O conjunto de dados utilizado possuía 1.782.000 registros do acelerômetro. O modelo de classificação foi construído com um algoritmo de Rede Neural Artificial utilizando 75% dos registros, enquanto os outros 25% foram utilizados na validação. Comparando a acurácia da classificação de cada classe, apresentada por

Nadimi *et al.* (2012), com as obtidas neste trabalho, constata-se que todas elas tiveram melhor resultado no modelo aqui proposto. O resultado é similar se comparado ao trabalho de Jesus (2014), que classificou o comportamento de vacas, utilizando o algoritmo *END* e os dados de posicionamento do animal.

As acurácias obtidas na validação dos modelos apresentados por Nadimi *et al.* (2012) e Jesus (2014) são apresentados na Tabela 39 e podem ser comparados com os resultados do modelo de classificação dos comportamentos primários proposto nesta pesquisa.

Tabela 39 – Comparativo do resultado da validação do modelo de classificação dos comportamentos primários com os apresentados por Nadimi *et al.* (2012) e Jesus (2014).

Classe	Acurácia		
	Nadimi <i>et al.</i> (2012)	Jesus (2014)	Esta pesquisa
Pastando/Procurando	83,8%	73,1%	96,4%
Andando	73,8%	87,5%	95%
Em Pé	71,8%	56,9%	88,6%
Deitado	83,2%	84,6%	93,4%

No trabalho de Dutta *et al.* (2015) foram utilizados sensores similares aos desta pesquisa (GPS + acelerômetro). O trabalho monitorou 24 vacas e obteve um conjunto de dados com 19.603 registros, dos quais 50% foram utilizados para treinar o modelo e os outros 50% para validá-lo, utilizando o algoritmo *Bagging* com um treinador baseado em árvore de decisão. O trabalho não diferenciou os momentos que o animal estava Em Pé e Deitado, apenas distinguindo os momentos que o animal estava Descansando e Ruminando, além dos momentos que estava Andando e Pastando.

Para possibilitar a comparação dos resultados, agrupou-se as classes Em Pé-Parado e Deitado-Parado na categoria Descansando e as classes Em Pé-Ruminando e Deitado-Ruminando na categoria Ruminando, e um novo modelo foi construído e validado, com o algoritmo *Random Forest* e o método *cross-validation*. A acurácia obtida foi de 89,99%, enquanto Dutta *et al.* (2015) obteve 96%. Na Tabela 41 pode-se comparar os resultados em cada classe.

Tabela 40 – Comparativo dos resultados obtidos com os apresentados por Dutta *et al.* (2015).

Classe	Acurácia	
	Dutta <i>et al.</i> (2015)	Esta pesquisa
Andando	97%	94,5%
Descansando	92	82,5%
Pastando	93%	96,4%
Ruminando	97%	83,8%

O trabalho de Dutta *et al.* (2015) só não obteve melhor resultado na classe Pastando. Uma diferença na técnica utilizada por eles foi analisar os dados em segmentos agrupados de 5 segundos, enquanto neste trabalho os dados foram analisados pontualmente, ou seja, apenas o momento em que haviam anotações do comportamento no *software* para observação. Além disso, eles capturaram os dados em frequência diferente:

4 Hz o GPS e 10 Hz o acelerômetro, enquanto aqui neste trabalho os sensores estavam configurados na frequência de 1 Hz.

O trabalho de [González *et al.* \(2015\)](#) também utilizou os dados capturados por um GPS e um acelerômetro, configurados na mesma frequência de [Dutta *et al.* \(2015\)](#), agrupados em segmentos de 10 segundos e divididos em dois subconjuntos. O algoritmo de classificação foi desenvolvido baseado em árvore de decisão. A Tabela 41 apresenta os resultados da validação do modelo, em cada um dos sub-conjuntos utilizados.

O comportamento Ruminando ficou com melhor resultado nos testes com os dois subconjuntos de [González *et al.* \(2015\)](#). Já a classe Andando tem melhor acurácia no modelo aqui proposto e as classes Descansando e Pastando possuem um subconjunto com melhor resultado para cada trabalho.

Tabela 41 – Comparativo dos resultados obtidos com os apresentados por [González *et al.* \(2015\)](#).

Classe	Acurácia		
	González <i>et al.</i> (2015)		Esta pesquisa
	Subconjunto A	Subconjunto B	
Andando	83,8%	24,7%	94,5%
Descansando	60,7%	86,3%	82,5%
Pastando	93,7%	98,4%	96,4%
Ruminando	96,9%	87,3%	83,8%

6.5 Considerações Finais

Neste capítulo foram apresentadas as técnicas utilizadas para processar e analisar os modelos de classificação dos comportamentos. Inicialmente os conjuntos de dados foram testados com uma série de algoritmos e analisados individualmente, avaliando o impacto que cada colar e animal utilizado no experimento em campo possui sobre o modelo de classificação gerado. Os resultados apontaram que não há grande impacto em relação ao animal ou ao colar utilizado.

Com os algoritmos e dados avaliados, apresentaram-se os resultados da validação dos modelos de classificação dos comportamentos primários, gerados com os algoritmos *Random Forest* e SVM. O modelo gerado no algoritmo *Random Forest* obteve o melhor resultado.

Em seguida, apresentou-se uma proposta para classificação de comportamentos secundários, mostrando o resultado da validação dos modelos para classificar os comportamentos Ruminando e Bebendo Água. Para o comportamento Ruminando, os resultados indicam que o melhor caso é incluí-lo como um comportamento primário, enquanto para Bebendo Água o modelo de classificação como comportamento secundário obteve melhores resultados (Seção 6.3.2.2).

Por fim, os resultados foram comparados a outros trabalhos que abordaram a

identificação automática do comportamento animal, a partir dos dados de GPS e acelerômetro. Os resultados das classificações aqui obtidos, em geral, são melhores que a maior parte dos trabalhos avaliados. Porém, na classificação do comportamento ruminando, os trabalhos que agruparam os dados em segmentos obtiveram melhor acurácia, o que indica que os resultados podem ser melhorados se tal estratégia for utilizada com os dados aqui utilizados.

7 Conclusão

A produção pecuária é uma das principais fontes econômicas do Brasil e a utilização de tecnologias na produção e na gestão dos negócios pecuários apresenta-se como uma poderosa ferramenta para estimular a competitividade e qualidade dos produtos. Este trabalho apresentou técnicas e ferramentas computacionais, no âmbito da Pecuária de Precisão, visando a identificação automática do comportamento bovino, a partir dos dados coletados por sensores e processados com algoritmos de reconhecimento de padrões.

Um equipamento, para coleta dos dados de movimentação e posicionamento do animal, foi projetado e quatro unidades dele construídas e validadas em campo. Os dados coletados pelo equipamento foram utilizados para construir um modelo de classificação do comportamento bovino. Os resultados da identificação automática dos comportamentos foram superiores aos apresentados por Jesus (2014), indicando que para aferir com acurácia os comportamentos é necessário o uso de sensores que capturem a movimentação do animal, considerando as características de movimentação de cada comportamento. O equipamento foi testado em campo por mais de 500 horas e apresentou-se adequado para a proposta. Alguns ajustes podem ser feitos para melhorar a sua apresentação e manuseio. Entre elas estão o encapsulamento do circuito, que foi adaptado em uma caixa plástica e pode ser acoplado em uma caixa específica projetada para tal, e alterar o método utilizado para prendê-lo ao animal, garantindo que o equipamento não se desloque com a movimentação do animal.

Para a predição dos comportamentos buscou-se avaliar os algoritmos de classificação supervisionada disponíveis no WEKA. Ao todo foram avaliados 57 algoritmos, disponíveis em 9 categorias, mas que não esgotam todos os disponíveis na ferramenta. Estes algoritmos exigem que os dados estejam rotulados com as classes que se deseja classificar. Considerando a dificuldade e os problemas para realizar tal tarefa manualmente, utilizou-se a abordagem e a ferramenta proposta por Jesus (2014). Entretanto, novas funcionalidades foram implementadas, atendendo os requisitos deste trabalho e permitindo que ela seja utilizada em outras atividades de observação do comportamento animal.

Baseado no trabalho de Kilgour *et al.* (2012), propôs-se duas abordagens para a classificação dos comportamentos, chamados de comportamentos primários e comportamentos secundários. O algoritmo selecionado, baseado na acurácia apresentada nos testes realizados, foi o *Random Forest* da categoria *trees* (Árvores de decisão). Na classificação dos comportamentos primários, a validação do modelo retornou uma acurácia média de 93,99%. Para classificação do comportamento Bebendo Água, considerado o modelo de classificação de comportamentos secundários, obteve-se uma acurácia de 94,1%. Na

classificação do comportamento Ruminando, apesar da proposta de classificação como comportamento secundário, o melhor resultado foi obtido quando ele foi classificado junto aos comportamentos primários, apontando acurácia média de 89,55%, atingindo 67,9% de acerto na classe Em Pé-Ruminando e 90,8% na classe Deitado-Ruminando.

Os resultados apontam a possibilidade de generalização do uso do equipamento e do modelo de classificação, para utilização em outras pesquisas, já que a validação se deu com dados de diferentes unidades dos equipamentos, capturados em diferentes animais de um mesmo rebanho. Porém, recomenda-se que novas pesquisas utilizem o equipamento e o modelo para aferir os comportamentos, em especial para animais de outros rebanhos e de outros pastos.

Por meio deste trabalho, buscou-se contribuir para as pesquisas na área de Pecuária de Precisão e para a parceria entre a Embrapa Gado de Corte e a FACOM/UFMS, apresentando ferramentas para a classificação automática dos comportamentos bovino.

7.1 Trabalhos Futuros

O colar utilizado para coleta dos dados dos animais não possui sistema de comunicação sem fio, exigindo que ele seja retirado do animal para que se acesse os dados coletados, interferindo na rotina natural do animal. Sugere-se a implementação de um sistema de comunicação sem fio, de forma que os dados coletados sejam enviados à uma estação base, a partir de RSSF. Assim, evita-se a interferência humana nos comportamentos do animal, exceto quando necessário a troca da bateria do equipamento. Vale ressaltar que o equipamento aqui desenvolvido possui suporte ao *XBee*, um rádio de comunicação sem fio baseado no protocolo *ZigBee* (padrão IEEE 802.15.4) e compatível com o Arduino Fio.

O foco dos trabalhos tem sido apenas identificar os comportamentos e indicar a localização do animal na área de pasto. Com os comportamentos e a localização identificados, é possível indicar como as condições do ambiente influenciam as reações do animal. Novos sensores podem aumentar o número de informações coletadas, incluindo as do ambiente, como temperatura, umidade, intensidade de sinais sonoros e luminosidade. Este último foi capturado neste trabalho, mas não foi utilizado durante o seu desenvolvimento.

Os dados foram coletados em intervalos de 1 segundo, considerando a necessidade de se formar uma base com grande volume de dados, buscando um modelo de classificação com ótima acurácia. Para um sistema de monitoramento do comportamento animal pode-se utilizar outro intervalo de coletas, baseado na experiência e orientação de especialistas em comportamento animal. Intervalos maiores indicam maior vida útil da bateria que alimenta o sistema e redução do número de vezes que o animal precisa ser recolhido para substituição da bateria.

Pode-se extratificar as coletas por tempo, a fim de identificar se é possível reduzir o número de coletas/observações para formar a base de dados para treinar um sistema de classificação do comportamento bovino. Pode-se ainda utilizar o equipamento para coletar os dados e avaliar a variação e frequência dos comportamentos e a partir dele indicar o melhor intervalo de coleta.

Para o monitoramento dos comportamentos sugere-se a implementação de um sistema que apresente informações detalhadas sobre os comportamentos classificados. O modelo gerado neste trabalho pode ser utilizado para tal, já que o WEKA pode ser incorporado ao código de um aplicativo Java. A aplicação pode apresentar: onde, quando e com qual frequência os comportamentos ocorreram. Estas informações podem auxiliar o gerenciamento do rebanho, apresentando suas características e possíveis variações nos padrões.

Apenas com os valores registrados pelos sensores pode não ser possível identificar todos os padrões de comportamentos, sendo necessário avaliar os padrões dos dados, em especial nos sensores de movimentação (SCHEIBE; GROMANN, 2006). Considerando que o comportamento Ruminando não foi classificado com boa acurácia, pode se aplicar a técnica de avaliar os padrões dos dados, em segmentos que agregem os dados de um período de coleta e mostre as variações da posição do pescoço do animal, diferente da abordagem aqui utilizada que utilizou os valores absolutos apresentados pelos sensores. Os trabalhos de González *et al.* (2015) e Dutta *et al.* (2015) utilizaram esta abordagem e avaliaram diferentes tamanhos de segmentos. Outra abordagem é gerar novos atributos, a partir daqueles capturados pelos sensores, como foi aplicado por Jesus (2014).

Por fim, a proposta de classificação de comportamentos com a combinação de variáveis ou com a análise de segmentos, podem ser utilizadas para aferir novos comportamentos, que indiquem variações nos comportamentos habituais do animal, como um animal doente ou com problemas com a mosca-dos-chifres.

Referências

- AKYILDIZ, I.; SU, W.; SANKARASUBRAMANIAM, Y.; CAYIRCI, E. Wireless sensor networks: a survey. *Computer Networks*, v. 38, n. 4, p. 393 – 422, 2002. ISSN 1389-1286. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1389128601003024>>. Citado 4 vezes nas páginas 27, 28, 29 e 30.
- ALTUN, K.; BARSHAN, B.; TUNÇEL, O. Comparative study on classifying human activities with miniature inertial and magnetic sensors. *Pattern Recognition*, v. 43, n. 10, p. 3605–3620, 2010. Disponível em: <<http://dblp.uni-trier.de/db/journals/pr/pr43-.html>>. Citado na página 31.
- ANDERSON, R. E. D.; CIBILS, A. Spatiotemporal cattle data—a plea for protocol standardization. *Positioning*, v. 4, n. 1, p. 115 – 136, 2013. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168159111004217>>. Citado 4 vezes nas páginas 34, 35, 36 e 37.
- ANDREOTTI, R. *Caracterização de Inibidores de Serinoproteases (BmTIS) Presentes em Larvas de Carrapatos Boophilus microplus e o seu Efeito no Controle da Infestação Parasitária em Bovinos*. Tese (Doutorado) — UNIFESP, São Paulo, 2002. Citado na página 22.
- ARDUINO. *Arduino Fio*. 2015. Disponível em: <<https://www.arduino.cc/en/Main/ArduinoBoardFio>>. Acesso em: 30 set. 2015. Citado 2 vezes nas páginas 11 e 59.
- BRASIL. *Produto Interno Bruto da agropecuária deve ser de R\$ 1,1 trilhão*. 2014. Disponível em: <<http://www.agricultura.gov.br/comunicacao/noticias/2014/12-/produto-interno-bruto-da-agropecuaria-deve-ser-de-rs-1-trilhao>>. Acesso em: 27 abr. 2015. Citado na página 21.
- BROOKS, C. B. C.; HARRIS, S. Effects of global positioning system collar weight on zebra behavior and location error. *The Journal of Wildlife Management*, v. 72, n. 2, p. 527–534, 2008. Citado na página 37.
- CARTHY, J. D. *O Estudo do Comportamento*. São Paulo: Editora Nacional, 1969. Citado na página 33.
- CLUTTON-BROCK, J. *A Natural History of Domesticated Animals*. London: Cambridge University Press, 1999. Citado na página 33.
- CÁCERES, E. N.; PISTORI, H.; TURINE, M. A. S.; PIRES, P. P.; SOARES, C. O.; CARROMEU, C. Computational precision livestock - position paper. In: *II Workshop of the Brazilian Institute for Web Science Research*. Rio de Janeiro: [s.n.], 2011. Citado na página 22.
- DIOSDADO, J. V.; BARKER, Z.; HODGES, H.; AMORY, J.; CROFT, D.; BELL, N.; CODLING, E. Classification of behaviour in housed dairy cows using an accelerometer-based activity monitoring system. *Animal Biotelemetry*, v. 3, n. 1, p. 15, 2015. ISSN

- 2050-3385. Disponível em: <<http://www.animalbiotelemetry.com/content/3/1/15>>. Citado 3 vezes nas páginas 52, 55 e 56.
- DUDA, R. O.; HART, P. E.; STORK, D. G. *Pattern Classification*. 2. ed. New York: Wiley-Interscience, 2000. Citado 4 vezes nas páginas 11, 38, 39 e 41.
- DUTTA, R.; SMITH, D.; RAWNSLEY, R.; BISHOP-HURLEY, G.; HILLS, J.; TIMMS, G.; HENRY, D. Dynamic cattle behavioural classification using supervised ensemble classifiers. *Computers and Electronics in Agriculture*, v. 111, p. 18 – 28, 2015. ISSN 0168-1699. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168169914003123>>. Citado 9 vezes nas páginas 14, 52, 53, 54, 55, 56, 102, 103 e 107.
- ELECTRONICS, P. R. . *MinIMU-9 v3 Gyro, Accelerometer, and Compass (L3GD20H and LSM303D Carrier)*. 2015. Disponível em: <<https://www.pololu.com/product/2468>>. Acesso em: 30 set. 2015. Citado 2 vezes nas páginas 11 e 58.
- ELECTRONICS, S. *SparkFun OpenLog - DEV-09530*. 2015. Disponível em: <<https://www.sparkfun.com/products/9530>>. Acesso em: 30 set. 2015. Citado 2 vezes nas páginas 11 e 58.
- ELECTRONICS, S. *SparkFun Venus GPS with SMA Connector - GPS-11058*. 2015. Disponível em: <<https://www.sparkfun.com/products/11058>>. Acesso em: 30 set. 2015. Citado 2 vezes nas páginas 11 e 57.
- EMBRAPA. *Embrapa - Missão, visão e valores*. 2015. Disponível em: <<https://www.embrapa.br/missao-visao-e-valores>>. Acesso em: 22 set. 2015. Citado na página 22.
- EMBRAPA. *Embrapa Gado de Corte - História*. 2015. Disponível em: <<https://www.embrapa.br/gado-de-corte/historia>>. Acesso em: 22 set. 2015. Citado na página 22.
- EMBRAPA. *Embrapa Gado de Corte - Missão, visão, valores*. 2015. Disponível em: <<https://www.embrapa.br/gado-de-corte/missao-visao-valores>>. Acesso em: 22 set. 2015. Citado na página 22.
- EMBRAPA GADO DE LEITE. *Panorama do Leite*. Ano 6. n. 65. Juiz de Fora, 2012. Citado na página 22.
- FERRAZ, M. R. *Manual do Comportamento Animal*. Rio de Janeiro: Editora Rubio, 2011. ISBN 978-85-7771-060-7. Citado na página 33.
- FILHO, K. E.; FIGUEIREDO, G. Retrospectiva e perspectivas de cruzamentos no Brasil. In: *Anais Simpósio Brasileiro Sobre Cruzamento de Bovinos de Corte*. Londrina: IAPAR, 2003. p. 11 – 35. CD-ROM. Citado na página 21.
- FRANCIS, J.; LITTLE, D. A. Resistance of droughtmaster cattle to tick infestation and babesiosis. *Australian Veterinary Journal*, v. 40, n. 7, p. 247–253, jul. 1964. Citado na página 22.
- FURLONG, J. *Carrapato: problemas e soluções*. Juiz de Fora: Embrapa Gado de Leite, 2005. ISBN 85-85748-63-X. Citado na página 22.

GONZÁLEZ, L.; BISHOP-HURLEY, G.; HANDCOCK, R.; CROSSMAN, C. Behavioral classification of data from collars containing motion sensors in grazing cattle. *Computers and Electronics in Agriculture*, v. 110, n. 0, p. 91 – 102, 2015. ISSN 0168-1699. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168169914002798>>. Citado 7 vezes nas páginas 14, 52, 53, 54, 55, 103 e 107.

GUO, Y.; POULTON, G.; CORKE, P.; BISHOP-HURLEY, G. J.; WARK, T.; SWAIN, D. L. Using accelerometer, high sample rate gps and magnetometer data to develop a cattle movement and behaviour model. *Ecological Modelling*, Elsevier, v. 220, n. 17, p. 2068 – 2075, 2009. Disponível em: <<http://eprints.qut.edu.au/33751/>>. Citado 3 vezes nas páginas 52, 53 e 56.

HALL, M.; FRANK, E.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P.; WITTEN, I. H. The weka data mining software: An update. *SIGKDD Explor. Newsl.*, ACM, New York, NY, USA, v. 11, n. 1, p. 10–18, nov. 2009. ISSN 1931-0145. Disponível em: <<http://doi.acm.org/10.1145/1656274.1656278>>. Citado na página 49.

HALL, S. J. Chillingham cattle: social and maintenance behaviour in an ungulate that breeds all year round. *Animal Behaviour*, v. 38, n. 2, p. 215 – 225, 1989. ISSN 0003-3472. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0003347289800843>>. Citado na página 34.

HANCOCK, J. Studies of grazing behavior in relation to grassland management i. variations in grazing habits of dairy cattle. *Agricultural Science*, v. 44, n. 4, p. 420–433, 1954. Citado na página 34.

HERBEL, C. H.; NELSON, A. Activities of hereford and santa gertrudis cattle on a southern new mexico range. *Journal of Range Management*, Society for Range Management, v. 19, p. 173–176, 1966. Disponível em: <[bibliography/098.pdf](#)>. Citado na página 34.

HSU, C.-W.; CHANG, C.-C.; LIN, C.-J. *A practical guide to support vector classification*. [S.l.], 2010. Disponível em: <<http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>>. Citado na página 92.

HUIRCÁN, J. I.; MUÑOZ, C.; YOUNG, H.; DOSSOW, L. V.; BUSTOS, J.; VIVALLO, G.; TONEATTI, M. Zigbee-based wireless sensor network localization for cattle monitoring in grazing fields. *Computers and Electronics in Agriculture*, v. 74, n. 2, p. 258 – 264, 2010. ISSN 0168-1699. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168169910001584>>. Citado na página 36.

HULBERT, I. A. R.; WYLLIE, J. T. B.; WATERHOUSE, A.; FRENCH, J.; MCNULTY, D. A note on the circadian rhythm and feeding behavior of sheep fitted with a lightweight gps collar. *Applied Animal Behavior Science*, v. 60, n. 4, p. 359–364, 1998. Citado na página 37.

IBGE. *Produção da Pecuária Municipal 2013*. 2013. Disponível em: <http://www.ibge.gov.br/home/estatistica/economia/ppm/2013/default_pdf.shtm>. Acesso em: 27 abr. 2015. Citado na página 21.

- IBGE. *Indicadores IBGE - Estatística da Produção Pecuária*. 2015. Disponível em: <http://www.ibge.gov.br/home/estatistica/indicadores/agropecuaria/producaoagropecuaria-abate-leite-couro-ovos_201401_publ_completa.pdf>. Acesso em: 27 abr. 2015. Citado na página 21.
- JESUS, L. de. *Identificação do Comportamento Bovino por meio do Monitoramento Animal*. Dissertação (Mestrado) — Universidade Federal de Mato Grosso do Sul, Campo Grande, 2014. Citado 16 vezes nas páginas 12, 14, 23, 36, 52, 54, 55, 56, 57, 70, 72, 75, 88, 102, 105 e 107.
- KILGOUR, R. J. In pursuit of “normal”: A review of the behaviour of cattle at pasture. *Applied Animal Behaviour Science*, v. 138, n. 1–2, p. 1 – 11, 2012. ISSN 0168-1591. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168159111004229>>. Citado 3 vezes nas páginas 33, 34 e 35.
- KILGOUR, R. J.; UETAKE, K.; ISHIWATA, T.; MELVILLE, G. J. The behaviour of beef cattle at pasture. *Applied Animal Behaviour Science*, v. 138, n. 1–2, p. 12 – 17, 2012. ISSN 0168-1591. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168159111004217>>. Citado 5 vezes nas páginas 33, 34, 35, 37 e 105.
- KRAUSE, J.; KRAUSE, S.; ARLINGHAUS, R.; PSORAKIS, I.; ROBERTS, S.; RUTZ, C. Reality mining of animal social systems. *Trends in Ecology Evolution*, v. 28, n. 9, p. 541 – 551, 2013. ISSN 0169-5347. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0169534713001468>>. Citado na página 52.
- LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. *Biometrics*, International Biometric Society, v. 33, n. 1, 1977. Citado 2 vezes nas páginas 13 e 47.
- LOUREIRO, A.; NOGUEIRA, J.; RUIZ, L.; MINI, R.; E.F.NAKAMURA; FIGUEIREDO, C. Redes de sensores sem fio. In: *Anais do XXI Simpósio Brasileiro de Redes de Computadores*. Natal: SBRC'2003, 2003. p. 179–226. Citado 5 vezes nas páginas 11, 27, 28, 29 e 30.
- MAPA. *Plano Agrícola e Pecuário 2013/2014 - Ministério da Agricultura, Pecuária e Abastecimento. Secretaria de Política Agrícola*. 2013. Disponível em: <http://www.agricultura.gov.br/arq_editor/file/acs/PAP20132014-web.pdf>. Acesso em: 27 abr. 2015. Citado na página 21.
- MAPA. *Brasil Projeções do Agronegócio 2013/2014 a 2023/2024 - Ministério da Agricultura, Pecuária e Abastecimento. Assessoria de Gestão Estratégica*. 2014. Disponível em: <<http://www.ceplac.gov.br/download/pa1324.pdf>>. Acesso em: 27 abr. 2015. Citado na página 21.
- MARTISKAINEN, P.; JÄRVINEN, M.; SKÖN, J.-P.; TIIRIKAINEN, J.; KOLEHMAINEN, M.; MONONEN, J. Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines. *Applied Animal Behaviour Science*, v. 119, n. 1–2, p. 32 – 38, 2009. ISSN 0168-1591. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168159109000951>>. Citado 7 vezes nas páginas 14, 23, 51, 52, 53, 55 e 101.

- MOREAU, M.; SIEBERT, S.; BUERKERT, A.; SCHLECHT, E. Use of a tri-axial accelerometer for automated recording and classification of goats' grazing behaviour. *Applied Animal Behaviour Science*, v. 119, n. 3–4, p. 158 – 170, 2009. ISSN 0168-1591. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168159109001324>>. Citado 5 vezes nas páginas 51, 52, 53, 54 e 55.
- NADIMI, E.; BLANES-VIDAL, V.; JØRGENSEN, R.; CHRISTENSEN, S. Energy generation for an ad hoc wireless sensor network-based monitoring system using animal head movement. *Computers and Electronics in Agriculture*, v. 75, n. 2, p. 238 – 242, 2011. ISSN 0168-1699. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168169910002322>>. Citado na página 36.
- NADIMI, E.; JØRGENSEN, R.; BLANES-VIDAL, V.; CHRISTENSEN, S. Monitoring and classifying animal behavior using zigbee-based mobile ad hoc wireless sensor networks and artificial neural networks. *Computers and Electronics in Agriculture*, v. 82, n. 0, p. 44 – 54, 2012. ISSN 0168-1699. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168169911003206>>. Citado 6 vezes nas páginas 14, 23, 36, 53, 101 e 102.
- NADIMI, E.; SØGAARD, H.; BAK, T. Zigbee-based wireless sensor networks for classifying the behaviour of a herd of animals using classification trees. *Biosystems Engineering*, v. 100, n. 2, p. 167 – 176, 2008. ISSN 1537-5110. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1537511008000834>>. Citado na página 36.
- NADIMI, E.; SØGAARD, H.; BAK, T.; OUDSHOORN, F. Zigbee-based wireless sensor networks for monitoring animal presence and pasture time in a strip of new grass. *Computers and Electronics in Agriculture*, v. 61, n. 2, p. 79 – 87, 2008. ISSN 0168-1699. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168169907002013>>. Citado 2 vezes nas páginas 36 e 53.
- OLIVEIRA, C. C. de. *Desempenho e Comportamento Ingestivo Diurno de Novilhas Nelore em Sistemas Integrados de Produção no Cerrado Brasileiro*. Dissertação (Mestrado) — Universidade Federal dos Vales do Jequitinhonha e Mucuri, Diamantina, 2013. Citado na página 37.
- OLIVEIRA, M. T. P. de. *Análise Comportamental de Bovinos Baseada em Trajetórias Semânticas Aplicada à Pecuária de Precisão*. Dissertação (Mestrado) — Universidade Federal de Mato Grosso do Sul, Campo Grande, 2013. Citado 3 vezes nas páginas 23, 52 e 54.
- OTIS, D. L.; WHITE, G. C. Autocorrelation of location estimates and the analysis of radiotracking data. *The Journal of Wildlife Management*, v. 63, n. 3, p. 1039–1044, 1999. Citado na página 34.
- OUDSHOORN, F. W.; KRISTENSEN, T.; NADIMI, E. S. Dairy cow defecation and urination frequency and spatial distribution in relation to time-limited grazing. *Live-stock Science*, v. 113, n. 1, p. 62–73, 2008. ISSN 1871-1413. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1871141307002168>>. Citado na página 34.

- ROBERT, B.; WHITE, B. J.; RENTER, D. G.; LARSON, R. L. Evaluation of three-dimensional accelerometers to monitor and classify behavior patterns in cattle. *Comput. Electron. Agric.*, Elsevier Science Publishers B. V., Amsterdam, The Netherlands, The Netherlands, v. 67, n. 1-2, p. 80–84, jun. 2009. ISSN 0168-1699. Disponível em: <<http://dx.doi.org/10.1016/j.compag.2009.03.002>>. Citado 4 vezes nas páginas 51, 52, 53 e 55.
- RUIZ-GARCIA, L.; LUNADEI, L. The role of RFID in agriculture: Applications, limitations and challenges. *Computers and Electronics in Agriculture*, v. 79, n. 1, p. 42 – 50, 2011. ISSN 0168-1699. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168169911001876>>. Citado na página 23.
- SCHEIBE, K.; GROMANN, C. Application testing of a new three-dimensional acceleration measuring system with wireless data transfer (WAS) for behavior analysis. *Behavior Research Methods*, Springer-Verlag, v. 38, n. 3, p. 427–433, 2006. ISSN 1554-351X. Disponível em: <<http://dx.doi.org/10.3758/BF03192796>>. Citado 9 vezes nas páginas 23, 31, 33, 36, 37, 51, 52, 53 e 107.
- SCHÜTZ, K. E.; ROGERS, A. R.; COX, N. R.; TUCKER, C. B. Dairy cows prefer shade that offers greater protection against solar radiation in summer: Shade use, behaviour, and body temperature. *Applied Animal Behaviour Science*, v. 116, n. 1, p. 28 – 34, 2009. ISSN 0168-1591. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168159108001883>>. Citado na página 37.
- SILVA, A. C. de S.; ARCE, A. I. C.; SOUTO, S.; COSTA, E. J. X. A wireless floating base sensor network for physiological responses of livestock. *Computers and Electronics in Agriculture*, v. 49, n. 2, p. 246 – 254, 2005. ISSN 0168-1699. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168169905000918>>. Citado na página 36.
- SILVEIRA, E. Maturidade precoce - cruzamentos, manejo e biologia molecular deixam bovinos prontos para o abate em 15 meses. *Revista Pesquisa - FAPESP*, Edição Especial de 50 Anos, p. 190 – 193, 2012. Disponível em: <<http://revistapesquisa.fapesp.br/2012/08/21/maturidade-precoce/>>. Citado na página 21.
- STMICROELECTRONICS. *MEMS motion sensor: three-axis digital output gyroscope - Datasheet - production data*. 2013. Disponível em: <<http://www.st.com/web/en/resource/technical/document/datasheet/DM00036465.pdf>>. Acesso em: 02 nov. 2015. Citado 2 vezes nas páginas 11 e 31.
- STMICROELECTRONICS. *Ultra-compact high-performance eCompass module: 3D accelerometer and 3D magnetometer - Datasheet - production data*. 2013. Disponível em: <<http://www.st.com/st-web-ui/static/active/en/resource/technical/document-datasheet/DM00027543.pdf>>. Acesso em: 02 nov. 2015. Citado 2 vezes nas páginas 11 e 31.
- TAN, P.-N.; STEINBACH, M.; KUMAR, V. *Introdução ao DATAMINING Mineração de Dados*. Rio de Janeiro, RJ, Brasil: Editora Ciência Moderna, 2009. ISBN 9788573937619. Citado 9 vezes nas páginas 11, 38, 39, 40, 41, 42, 43, 44 e 46.
- TAOS. *TSL2560, TSL2561 Light-to-Digital Converter*. 2005. Disponível em: <<http://cdn.sparkfun.com/datasheets/Sensors/LightImaging/TSL2561.pdf>>. Acesso em: 02 nov. 2015. Citado na página 33.

- TECHNOLOGY, S. *Venus638FLPx GPS Receiver - Data Sheet*. 2008. Disponível em: <http://cdn.sparkfun.com/datasheets/Sensors/GPS/Venus/638/doc/Venus638FLPx_DS_v07.pdf>. Acesso em: 02 nov. 2015. Citado na página 79.
- THEODORIDIS, S.; KOUTROUMBAS, K. *Pattern Recognition*. 3. ed. San Diego: Academic Press, 2006. Citado na página 37.
- WATANABE, S.; IZAWA, M.; KATO, A.; ROPERT-COUDERT, Y.; NAITO, Y. A new technique for monitoring the detailed behaviour of terrestrial animals: A case study with the domestic cat. *Applied Animal Behaviour Science*, v. 94, n. 1–2, p. 117 – 131, 2005. ISSN 0168-1591. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0168159105000353>>. Citado 7 vezes nas páginas 31, 32, 35, 37, 51, 52 e 55.
- WEBTRONICO. *LDR 5mm - Sensor de luminosidade*. 2015. Disponível em: <<http://www.webtronico.com/ldr-5mm-sensor-de-luminosidade.html>>. Acesso em: 30 set. 2015. Citado 2 vezes nas páginas 11 e 58.
- WEKA. *Machine Learning Group*. 2015. Disponível em: <<http://www.cs.waikato.ac.nz/ml/index.html>>. Acesso em: 08 ago. 2015. Citado na página 47.
- ZEMO, T.; KLEMMEDSON, J. Behavior of fistulated steers on a desert grassland. *Journal of Range Management Archives*, v. 23, n. 3, p. 158–163, 1970. Citado na página 34.