

Alex Sandro de Paula Rodrigues

**Tecnologia RFID aplicada à Agrocomputação:
um estudo de caso utilizando descoberta de
conhecimento em base de dados**

Rio Grande, Brasil.

16 de março de 2017.

Alex Sandro de Paula Rodrigues

Tecnologia RFID aplicada à Agrocomputação: um estudo de caso utilizando descoberta de conhecimento em base de dados

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Computação da Universidade Federal do Rio Grande, como requisito para a obtenção do grau de Mestre em Engenharia de Computação

Universidade Federal do Rio Grande
Mestrado em Engenharia de Computação

Orientador: Prof. Dra. Regina Barwaldt

Rio Grande, Brasil.
16 de março de 2017.

Ficha catalográfica

R696t Rodrigues, Alex Sandro de Paula.
Tecnologia RFID aplicada à Agrocomputação: um estudo de caso utilizando descoberta de conhecimento em base de dados / Alex Sandro de Paula Rodrigues. – 2017.
65 p.

Dissertação (mestrado) – Universidade Federal do Rio Grande – FURG, Programa de Pós-graduação em Computação, Rio Grande/RS, 2017.
Orientadora: Dr^a. Regina Barwaldt.

1. Avicultura 2. RFID 3. Descoberta de conhecimento em base de dados I. Barwaldt, Regina II. Título.

CDU 004.65:636.5



UNIVERSIDADE FEDERAL DO RIO GRANDE
Centro de Ciências Computacionais
Programa da Pós-Graduação em Computação
Curso de Mestrado em Engenharia de Computação

ATA DE SESSÃO DE DEFESA DE DISSERTAÇÃO DE MESTRADO

Ata No. 03/2017

Na data de 16 de março de 2017, às 9 horas, ocorreu a Sessão de Defesa de Dissertação de Mestrado de Engenharia de Computação, que apresentou a dissertação intitulada **Tecnologia RFID aplicada à Agrocomputação: um estudo de caso utilizando descoberta de conhecimento em base de dados**, realizada sob a orientação da Profa. Dra. Regina Barwaldt. A banca examinadora foi constituída pelos Profs. Nelson José Laurino Dionello (UFPEL), Eduardo Nunes Borges (FURG) e Vagner Santos da Rosa (FURG), sob a presidência do orientador. Após a apresentação do trabalho, a banca arguiu o candidato e, a seguir, deliberou pela

- aprovação da Dissertação
- aprovação da Dissertação, sugerindo modificações no texto
- reprovação da Dissertação

Rio Grande, 16 de março de 2017.

Prof. Dr. Nelson José Laurino Dionello

Prof. Dr. Eduardo Nunes Borges

Prof. Dr. Vagner Santos da Rosa

Prof. Dra. Regina Barwaldt
Orientador(a)

UNIVERSIDADE FEDERAL DO RIO GRANDE
Centro de Ciências Computacionais
Programa de Pós-Graduação em Engenharia de Computação
Curso de Mestrado em Engenharia de Computação

DISSERTAÇÃO DE MESTRADO

**Tecnologia RFID aplicada à Agrocomputação: um estudo de caso utilizando
descoberta em conhecimento de dados**

Banca Examinadora:

Profa. Dra. Regina Barwaldt
Orientadora

Prof. Dr. Nelson José Laurino Dionello

Prof. Dr. Eduardo Nunes Borges

Prof. Dr. Vagner Santos da Rosa

Agradecimentos

DEDICO esta conquista aos meus familiares, que, constantemente, me desafiam a mobilizar o melhor de mim, constituindo-se como referências em minha vida pessoal e profissional.

AGRADEÇO:

- a Deus pela vida e pelas possibilidades de múltiplos encontros com pessoas que contribuíram para a realização deste trabalho;
- Universidade Federal do Rio Grande, pela oportunidade de aprofundar meus conhecimentos e aprimorar meu desempenho profissional;
- Escola Técnica Estadual Santa Isabel pelo espaço cedido para o desenvolvimento do presente trabalho, assim como, ao Prof. Gilberto Barwaldt pelo apoio;
- a Professora Dra. Regina Barwaldt, por ter acreditado em meu potencial e me incentivado a vencer os desafios desta trajetória acadêmica;
- aos professores e colegas do Curso de Mestrado em Engenharia de Computação, pelas valiosas contribuições, bem como, o enriquecedor intercâmbio de ideias e pela disponibilidade e inestimável colaboração na pesquisa;
- Agradeço minha Sastria de Paula Rodrigues, que me deu apoio e incentivo nos momentos de desânimo.
- aos amigos, que, na convivência fraterna, se tornaram apoio indispensável à realização deste trabalho.

Resumo

Introdução: Avicultura é um importante setor da economia brasileira, representando 1.5% do PIB brasileiro, bem como, cinco milhões de empregos diretos e indiretos. Destaca-se também a privilegiada posição brasileira no cenário mundial, estando entre os três primeiros do mundo, quando trata-se da produção e consumo da ave, e em primeiro lugar quando refere-se a exportações da mesma.

Objetivo: Como forma de contribuir com este contexto, o presente trabalho buscou determinar por intermédio da descoberta de conhecimento em base de dados um modelo de classificação, apoiado em variáveis ambientais, como: temperatura, umidade relativa do ar, luminosidade, peso e idade dos frangos de corte, capaz de prever o peso dos frangos de corte, perante as variáveis citadas previamente.

Metodologia: Para efetivar tal tarefa, foi utilizado o método de monitoramento automático, RFID, que pode ser visto como uma tecnologia de identificação que opera sob ondas de rádio, capaz de enviar e armazenar dados de forma remota. Concomitante ao uso desta tecnologia, foi empregada a descoberta de conhecimento em bases de dados com objetivo de descobrir padrões implícitos no conjuntos de dados, transformando os dados sem significado em informação útil.

Conclusão: Por fim, o presente trabalho apresenta dois estudos de caso, mostrando como as variáveis citadas anteriormente, influenciaram no desempenho comportamental e nutricional dos frangos de corte, bem como, o modelo classificatório citado anteriormente. Por meio do modelo, foi possível compreender a atuação das variáveis climáticas no processo de ganho de peso das aves de corte.

Palavras-chave: Avicultura; RFID; Descoberta de conhecimento em base de dados.

Abstract

Introduction: Poultry is an important sector of the Brazilian economy, accounting for 1.5 % of Brazil's GDP, as well as five million direct and indirect jobs. Also noteworthy is the privileged Brazilian position on the world stage, being among the top three in the world when it is the production and consumption of the bird, and first when it comes to exports of the same.

Objective: As a way to contribute to this context, the present work sought to develop, through the discovery of knowledge in a database, a mathematical model, supported by environmental variables such as: temperature, relative air humidity, lightness, weight and age of broilers , Capable of predicting the weight of the broilers, in relation to the variables previously mentioned.

Methodology: To accomplish this task, an automatic tracking method, RFID, which can be seen as an identification technology that operates on radio waves capable of sending and storing data remotely. Concomitant to the use of this technology, will be used to knowledge discovery in databases in order to discover implicit patterns in data sets, transforming data into useful information meaningfully.

Conclusion: Finally, the present work presents two case studies, showing as variables previously mentioned, influenced on the behavioral and nutritional performance of broilers, as well as the classificatory model mentioned above. Through the model, it was possible to understand the performance of climatic variables in the process of weight gain of broilers.

Keywords: Aviculture; RFID; Knowledge Discovery in Database.

Lista de ilustrações

Figura 1 – Produção da carne de frango	14
Figura 2 – Exportação da carne de frango	15
Figura 3 – Consumo da carne de frango	15
Figura 4 – Antena de polarização linear	17
Figura 5 – Antena de polarização circular	17
Figura 6 – Funcionamento do sistema RFID	18
Figura 7 – Etapas do processo de descoberta de conhecimento em base de dados	19
Figura 8 – Razão estatística	23
Figura 9 – Razão computacional	23
Figura 10 – Razão representacional	24
Figura 11 – Modelo entidade-relacionamento	35
Figura 12 – Local disponibilizado para coleta de dados	37
Figura 13 – <i>Tag</i> RFID no frango	38
Figura 14 – Sensores usados no desenvolvimento da proposta	38
Figura 15 – Fonte de calor externa	43
Figura 16 – Ambiente de criação dos frangos	44
Figura 17 – Posicionamento <i>tag</i> no estudo de caso 2	44
Figura 18 – Resultado do algoritmo M5P	50
Figura 19 – Gráfico de procura por alimentação por períodos do dia	53
Figura 20 – Gráfico de evolução com relação a procura por alimentação	54
Figura 21 – Gráfico comparativo dos frangos de corte com relação a busca por ali- mentação	54
Figura 22 – Ambiente de criação dos frangos	55
Figura 23 – Leitura RFID x Sensores Temperatura	56
Figura 24 – Leitura RFID x Sensores Umidade	57
Figura 25 – Leitura RFID x Sensores Luminosidade	57

Lista de tabelas

Tabela 1 – Repositório selecionados para pesquisa	26
Tabela 2 – Palavras-chave escolhidas para o processo de busca	27
Tabela 3 – Consulta inicial de trabalhos utilizando termos em português	27
Tabela 4 – Consulta inicial de trabalhos utilizando termos em inglês	28
Tabela 5 – Consulta a partir da combinação de dois termos em língua portuguesa	28
Tabela 6 – Consulta a partir da combinação de dois termos em língua inglesa	29
Tabela 7 – Consulta a partir da combinação de três termos em língua portuguesa	29
Tabela 8 – Consulta a partir da combinação de três termos em língua inglesa	30
Tabela 9 – Consulta a partir da combinação de quatro termos em língua portuguesa	30
Tabela 10 – Consulta a partir da combinação de quatro termos em língua inglesa	30
Tabela 11 – Trabalhos relacionados escolhidos após a etapa de busca	31
Tabela 12 – Equipamentos utilizados nos estudos de caso	37
Tabela 13 – Tabela antes do pré-processamento dos dados (Dados Fictícios)	39
Tabela 14 – Tabela após o pré-processamento dos dados (Dados Fictícios)	40
Tabela 15 – Primeira coleta de peso	45
Tabela 16 – Segunda coleta de peso	45
Tabela 17 – Terceira coleta de peso	46
Tabela 18 – Quarta coleta de peso	46
Tabela 19 – Tabela antes do pré-processamento dos dados	47
Tabela 20 – Intervalos de temperatura e umidade - (TINOCO, 1988) e (ABREU; ABREU, 2005)	48
Tabela 21 – Tabela após o pré-processamento dos dados	49
Tabela 22 – Conversão alimentar	55

Lista de abreviaturas e siglas

CSS – *Cascading style sheets*

ETESI – Escola Técnica Estadual Santa Isabel

FURG – Universidade Federal do Rio Grande

HTML – Hypertext markup language

KDD – *Knowledge Discovery in Databases*

KG - Quilo

MER – Modelo entidade-relacionamento

PHP – *Hypertext Preprocessor*

PIB – Produto Interno Bruto

RFID – *Radio frequency identification*

SGBD – Sistema gerenciador de banco de dados

SQL - *Structured query language*

Sumário

1	INTRODUÇÃO	11
1.1	Justificativa	12
1.2	Objetivo Geral	12
1.2.1	Objetivos Específicos	12
1.3	Estrutura do Trabalho	13
2	REFERENCIAL TEÓRICO	14
2.1	Avicultura	14
2.2	RFID	16
2.2.1	Componentes do RFID	16
2.2.2	Funcionamento do sistema RFID	18
2.3	Descoberta de Conhecimento em Base de Dados	18
2.3.1	Tarefas Preditivas	20
2.3.2	Tarefas Descritivas	21
2.4	Sabedoria das Massas	22
2.4.1	Justificativas para o uso do <i>Ensemble Learning</i>	22
2.4.2	Construção de uma solução com <i>ensemble learning</i>	24
3	Trabalhos Relacionados	26
3.1	Definições e triagens	26
3.2	<i>Temperature management for the quality assurance of a perishable food supply chain</i>	31
3.3	<i>Wireless sensor networks (WSNs) in the agricultural and food industries</i>	32
3.4	Tecnologias da Informação e Comunicação e suas relações com a agricultura	33
4	METODOLOGIA	34
4.1	Tecnologias e linguagens utilizadas no desenvolvimento do trabalho	34
4.1.1	SQL - <i>Structured Query Language</i>	34
4.1.2	PostgreSQL	34
4.1.3	Arduino	34
4.1.4	Weka	35
4.2	Modelo entidade-relacionamento (MER)	35
4.3	Coleta de dados	36
4.4	Estudo de caso 1	39
4.4.1	Pré-processamento - Estudo de caso 1	39
4.4.2	Mineração de dados - Estudo de caso 1	41

4.4.3	Pós-processamento - Estudo de caso 1	42
4.5	Estudo de caso 2	43
4.5.1	Descoberta de Conhecimento em Base de Dados - Estudo de Caso 2	47
4.5.2	Pré-processamento - Estudo de caso 2	47
4.5.3	Mineração de dados - Estudo de caso 2	49
4.5.4	Pós-processamento - Estudo de caso 2	50
5	CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS	59
	Referências	61
	Apêndices	63
	APÊNDICE A Sistema de Informação	64

1 INTRODUÇÃO

O meio rural tem sofrido constantes transformações, seja no modo de produção e/ou distribuição, bem como, em outros integrantes do meio. Um dos aspectos que vêm contribuindo para essas mudanças é o uso crescente de de tecnologias, buscando agregar mais qualidade nos processos e melhoramento na produtividade. O agronegócio é um setor estratégico para economia brasileira, pois representa cerca de 23% do PIB (Produto Interno Bruto) brasileiro e 40% das exportações feitas pelo país. Neste meio destaca-se a avicultura uma das culturas mais representativas da pecuária brasileira, na qual, corresponde a 1,5% do PIB, 5 milhões de empregos diretos e indiretos, bem como, 47% da carne produzida no Brasil. Outro dado significativo acerca da avicultura brasileira é o seu desempenho no mercado internacional, onde o Brasil é o maior exportador de carne de aves do mundo e segundo maior produtor da mesma (MENDES, 2014), (AGRICULTURA, 2015). Ressalta-se ainda as projeções do Ministério da Agricultura, a respeito da avicultura entre os períodos de 2014/15 a 2024/25, onde é previsto um crescimento em termos de consumo e produção, a redor de 3% ao ano.

Destaca-se também que a motivação para presente pesquisa surgiu, a partir, de demandas pendentes do trabalho de conclusão de curso denominado *Inserção de tecnologia no processo de monitoramento em cadeia produtiva avícola* de (RODRIGUES, 2014), onde buscou-se, por meio, de tecnologia de monitoramento e análise de dados, entender como os períodos do dia e a idade das aves influenciavam no processo de busca por alimentação das mesmas. Neste mesmo trabalho, também houve o desenvolvimento de um sistema de informação (apêndice A) direcionado para gestão de aviários.

Dessa forma, o presente trabalho procura dar continuidade ao estudo anterior, apresentando uma nova metodologia de coleta e análise de dados sobre o comportamento de frangos de corte na região rural do município de São Lourenço do Sul. Essa análise busca estabelecer um modelo de classificação, capaz de prever o peso das aves considerando variáveis, como: temperatura, umidade relativa do ar e luminosidade, bem como, peso e a idade dos frangos de corte. Buscando assim compreender, como as mesmas, influenciam no desempenho nutricional dos frangos.

Para tal fim, adotou-se a tecnologia RFID (do inglês *Radio Frequency Identification*), um método de identificação automática através de sinais de rádio. Conjuntamente, a tecnologia de monitoramento, é empregado o KDD (do inglês *Knowledge Discovery in Databases*) cujo objetivo é descobrir padrões implícitos em conjuntos de dados, transformando os dados sem significado em informação útil. Por fim, o presente trabalho procura verificar se o uso do *ensemble learning*, que pode ser explicado, como um método de ob-

tenção de informação, onde um conjunto de algoritmos fornece respostas individuais que de alguma forma são reunidas, fornecendo uma resposta final, traz uma maior precisão na informação conquistada (DOMDOUZIS; KUMAR; ANUMBA, 2007), (FAYYAD et al., 1996), (KUNCHEVA, 2004).

1.1 Justificativa

A implantação de ferramentas de monitoramento e análise de dados neste cenário, é justificada devido aos benefícios que as mesmas oferecem ao produtor, visto que, por meio delas é possível obter um maior conhecimento sobre o comportamento dos frangos de corte, bem como, compreender como as variáveis ambientais impactam no processo de crescimento e conversão alimentar das aves de corte. Ressaltasse que a adoção de tecnologias no campo tende a aumentar o índice de produtividade, assim como, a eficácia no manejo dos insumos, produzindo assim, redução de custos com mão de obra e aumento na qualidade de trabalho.

Por fim, se deve salientar o caráter extensionista do presente trabalho, visto que, que o mesmo tem importante relevância, no que tange a aproximação do ensino superior com o ensino agrotécnico, bem como, o aperfeiçoamento desta área de pesquisa na região e arredores, onde há concentração de produtores rurais, engajados através das cooperativas.

1.2 Objetivo Geral

O objetivo geral do presente trabalho é propor um modelo para previsão do peso de frangos de corte, por meio de algoritmos de descoberta de conhecimento em base de dados, fundamentado em variáveis do ambiente, como: temperatura, umidade relativa do ar e luminosidade, da mesma forma que, foram considerados dados relativos a idade e peso das aves.

1.2.1 Objetivos Específicos

- Realizar revisão sistemática sobre o uso de RFID, descoberta de conhecimento em base de dados no âmbito da avicultura;
- Definir equipamentos RFID, bem como, sensores de supervisionamento ambiental a serem utilizados no estudo;
- Testar os equipamentos RFID em ambiente fechado, com a intenção de simular as condições da Escola Estadual Santa Isabel;

- Testar os sensores de temperatura, umidade e luminosidade em ambiente controlado com propósito de analisar a precisão das leituras;
- Implantar o sistema RFID no local de estudo, a fim de, coletar dados acerca dos frangos de corte;
- Implantar sensores de temperatura, umidade e luminosidade no ambiente de criação das aves, com objetivo de acompanhar suas condições;
- Modelagem e implantação do sistema para persistência dos dados coletados pelo RFID e sensores Arduino;
- Testar diferentes algoritmos de descoberta de conhecimento, buscando analisar quais apresentam melhor desempenho para proposta do trabalho;
- Realização de estudos de caso com os dados coletados para geração de informações pertinentes aos produtores de frango de corte;
- Avaliação dos modelos gerados pelo algoritmos de KDD durante o processo de análise de dados;

1.3 Estrutura do Trabalho

Este trabalho consiste em cinco capítulos, sendo que o primeiro e segundo são de cunho teórico relacionado ao tema, o terceiro apresenta a metodologia de pesquisa, bem como, os trabalhos relacionados considerados relevantes a presente proposta, o quarto apresenta as tecnologias e métodos para adotados atingir o objetivo específico do trabalho. Por fim, quinto capítulo aponta as considerações finais e trabalhos futuros do presente trabalho.

2 REFERENCIAL TEÓRICO

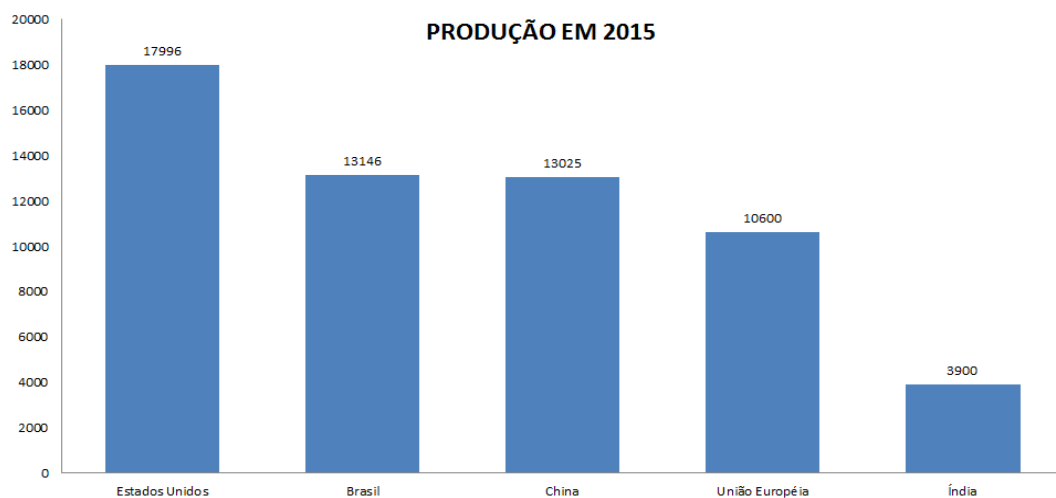
2.1 Avicultura

A avicultura é o ramo que destina-se à criação de aves e é uma das atividades mais importantes do agronegócio brasileiro, tendo seu início no século XIV, quando navegadores portugueses trouxeram os primeiros espécimes de frango para o Brasil. Em termos comerciais, a mesma desenvolveu-se a partir da década de 50, na região sudeste, onde aviários, associações e cooperativas foram estruturadas, assim como, novos métodos de manejo foram desenvolvidos (GIROTTI; AVILA, 2003), (SORJ; POMPERMAYER; CORADINI, 2008).

Ainda de acordo com os autores, a expansão do setor avícola em termos de mercado, aconteceu nas décadas de 60 e 70, devido à crescente importância do produto na dieta do brasileiro, bem como, a implantação do decreto nº 55.981, de 22 de abril de 1965 que surgiu para regulamentar a importação de aves de reprodução no Brasil. A partir, de tal decreto o país interrompeu a compra de matrizes oriundas de outros países e passou a produzi-las internamente.

Na atualidade a avicultura brasileira corresponde a 1,5% do PIB e 47% da carne produzida no Brasil conforme dados de (MENDES, 2014), (AVICULTURA, 2016). Na figura 1 pode-se observar os números da avicultura brasileira em termos de produção, no qual, em 2015 o país produziu mais de 13 milhões de toneladas de carne de frango, colocando o Brasil entre os três maiores produtores da mesma no mundo.

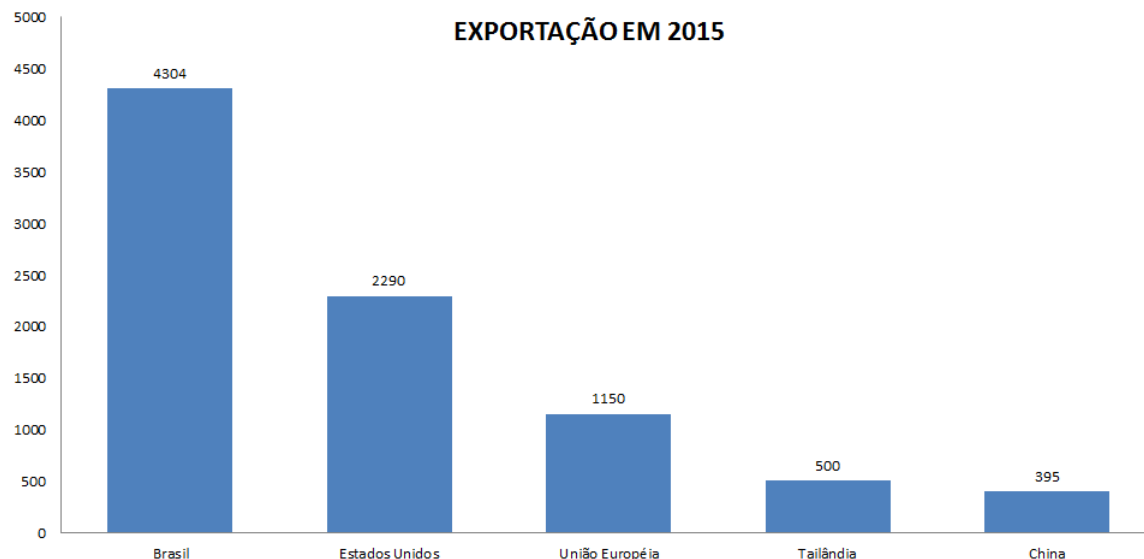
Figura 1: Produção da carne de frango



Fonte: Adaptado de (AVICULTURA, 2016)

Com relação as exportações, é possível observar na figura 2, que o país é o maior exportador da carne de frango, a frente de países, como: Estados Unidos e China.

Figura 2: Exportação da carne de frango

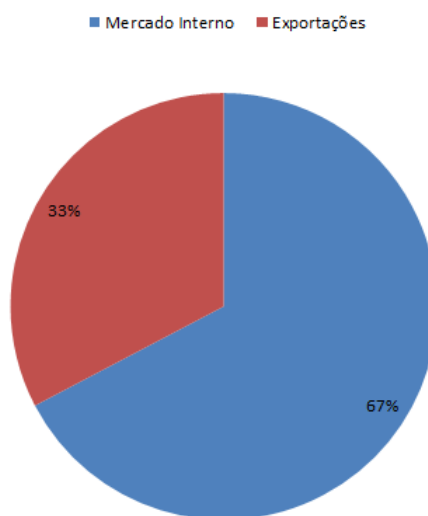


Fonte: Adaptado de (AVICULTURA, 2016)

Outro aspecto que demonstra a importância do setor avícola para o país, está no total produzido em 2015 (figura 3), quando 67% da carne foi destinada para consumo interno e o restante para exportações.

Figura 3: Consumo da carne de frango

CONSUMO DA CARNE DE FRANGO EM 2015



Fonte: Adaptado de (AVICULTURA, 2016)

Ressalta-se ainda as projeções realizadas pelo Ministério da Agricultura para o agronegócio brasileiro entre os anos de 2014/15 a 2024/25, onde, de acordo com o órgão federal esta previsto um crescimento em termos de consumo e produção.

Entre as culturas destacadas no relatório encontram-se a bovina, suína e avícola com expectativa de crescimento na produção em torno de 3%. Além do aumento com relação a produção também é projetada uma elevação em termos de consumo, principalmente da carne de frango, onde espera-se um crescimento médio anual de 2.8%. Por fim, o relatório demonstra grande confiança com relação as exportações, com destaque a cultura avícola com expansão anual esperada de 3,6% ao ano (MENDES, 2014), (AVICULTURA, 2016), (AGRICULTURA, 2015).

Finalizando, a partir dos dados mostrados, pode-se inferir a importância do setor de avicultura para a economia brasileira .

2.2 RFID

O uso de tecnologias de identificação tem-se destacado devido a sua capacidade de garantir um eficiente e efetivo sistema de gerenciamento, como também, fornecer dados confiáveis e precisos a respeito do que está sendo monitorado. Um exemplo é o RFID (*Radio Frequency Identification*) que pode ser entendido como uma tecnologia que opera sob ondas rádio, capaz de enviar e armazenar dados de forma remota, além de identificar, localizar e monitorar a posição de pessoas, animais ou outros objetos (GOMES, 2007).

2.2.1 Componentes do RFID

Um sistema RFID é composto por três elementos, sendo eles: *tags*, leitores e antenas, cada qual, com uma determinada atribuição na estrutura. O primeiro componente citado, a *tag* tem a função de identificar os diferentes elementos, nos quais, encontra-se acoplada. Estes identificadores diferencia-se em dois tipos: *tags* passivas ou ativas. As etiquetas passivas necessitam da energia gerada pelo campo eletromagnético do leitor para funcionarem, já as *tags* ativas possuem uma fonte de energia própria, o que elimina a necessidade de energizar o chip através da energia do campo eletromagnético (DOMDOUZIS; KUMAR; ANUMBA, 2007).

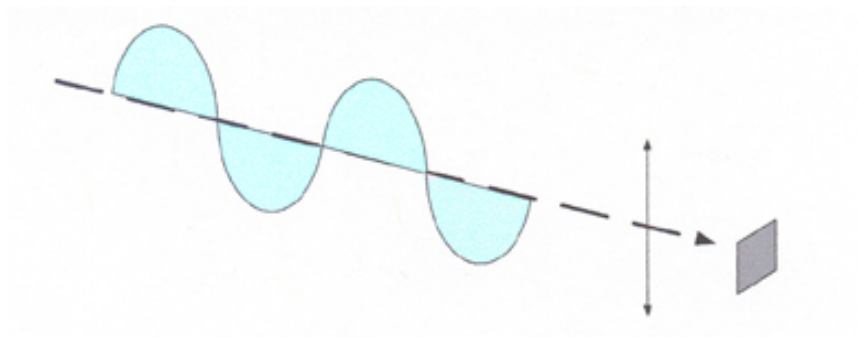
O segundo elemento do sistema RFID, o leitor tem a função de compreender as informações enviadas pelo identificador e encaminhá-las para sistemas externos de processamento de dados. Determinados autores consideram o leitor como o principal componente do sistema, visto que, ele tem a responsabilidade de controlar acessos múltiplos, realizar a correção de erros e entre outros. De acordo com (GOMES, 2007),

a razão da grande maioria destes processamentos serem colocados no

reader, advém do facto de o tag ser um dispositivo de tamanho reduzido e baixa complexidade (baixo custo), pelo que todos os mecanismos de segurança, gestão e controlo do sistema deverão ser colocados no reader [...]

O último componente do RFID, a antena é responsável por estabelecer a comunicação entre a *tag* e o leitor, além de, obter informações contidas na mesma. A emissão de sinais pela antena pode ocorrer de duas formas, por polarização linear ou circular. O funcionamento dos dois tipos é similar, diferenciando-se basicamente quanto ao modo de leitura das informações. Nas antenas de polarização linear (figura 4), as ondas eletromagnéticas se propagam em um só plano, podendo ele ser vertical ou horizontal.

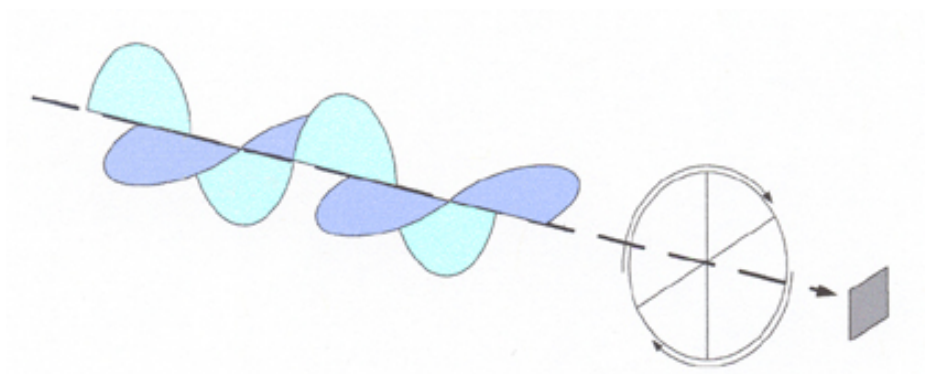
Figura 4: Antena de polarização linear



Fonte: (RODRIGUES, 2014)

Nas antenas de polarização circular (figura 5), as ondas eletromagnéticas se propagam em dois sentidos, criando um efeito circular, possibilitando a leitura das *tags* em qualquer sentido (DOMDOUZIS; KUMAR; ANUMBA, 2007).

Figura 5: Antena de polarização circular

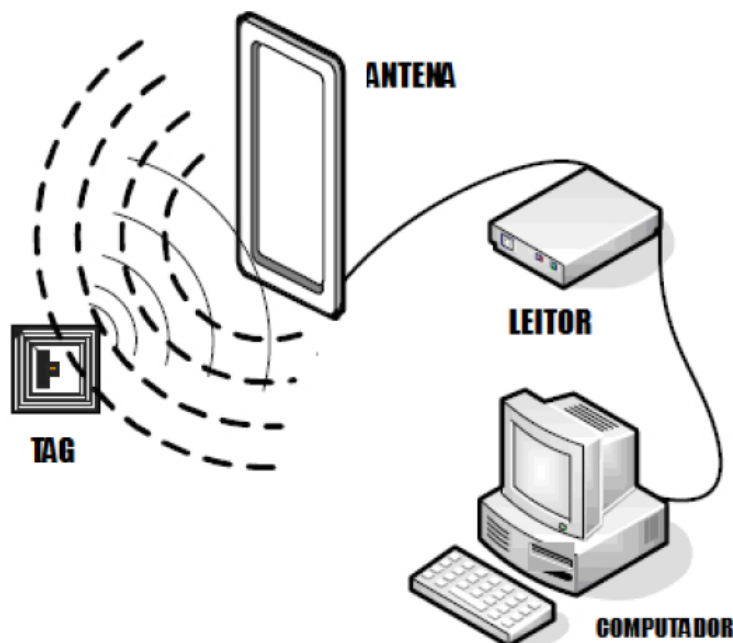


Fonte: (RODRIGUES, 2014)

2.2.2 Funcionamento do sistema RFID

Conforme mencionado anteriormente, um sistema de RFID é composto por: *tag*, leitor e antena. Na figura 6, está representado o funcionamento de um sistema RFID.

Figura 6: Funcionamento do sistema RFID



Fonte: Adaptada de (DOMDOUZIS; KUMAR; ANUMBA, 2007)

A comunicação do sistema RFID acontece por ondas de rádio que carregam informação. Usualmente, quando a antena percebe a *tag* em seu campo de leitura, transmite a informação, contida no dispositivo, para o leitor, que por sua vez, converte as ondas de rádio do RFID para informações digitais. Após a transformação os dados coletados poderão ser lidos, interpretados e analisados por sistemas processamento de dados (BARWALDT et al., 2014).

2.3 Descoberta de Conhecimento em Base de Dados

De acordo com (SILVA, 2004), atualmente vivemos num cenário onde pessoas e empresas têm produzido e acumulado grandes quantidades de dados, seja através de transações comerciais e/ou por comunicação. A grande disponibilidade de dados gera a eminente necessidade de transformá-los em informações e conhecimento úteis, capazes de auxiliar no processo de tomada de decisão. Isso tem exigido grandes investimentos da comunidade acadêmica, assim como, do setor privado que buscam na tecnologia agregar valor, agilizar operações e viabilizar inovações, visto que, a informação e conhecimento conquistados podem ser usados para diversos propósitos.

Conforme (SILVA, 2004), (TAN; STEINBACH; KUMAR, 2009), (WITTEN; FRANK, 2005) as ferramentas e técnicas empregadas para análise destes repositórios são os objetos estudo tratados pelo campo emergente da descoberta de conhecimento em bancos de dados (DCBD), da expressão em inglês *Knowledge Discovery in Databases* (KDD), que surgiu devido ao fato de ferramentas e técnicas tradicionais de análise de dados serem ineficazes com relação à avaliação de grandes volumes de dados, além das características não triviais que alguns conjuntos de dados apresentavam, não sendo possível com isso, a aplicação das técnicas existentes, tornando indispensável a criação de novos métodos.

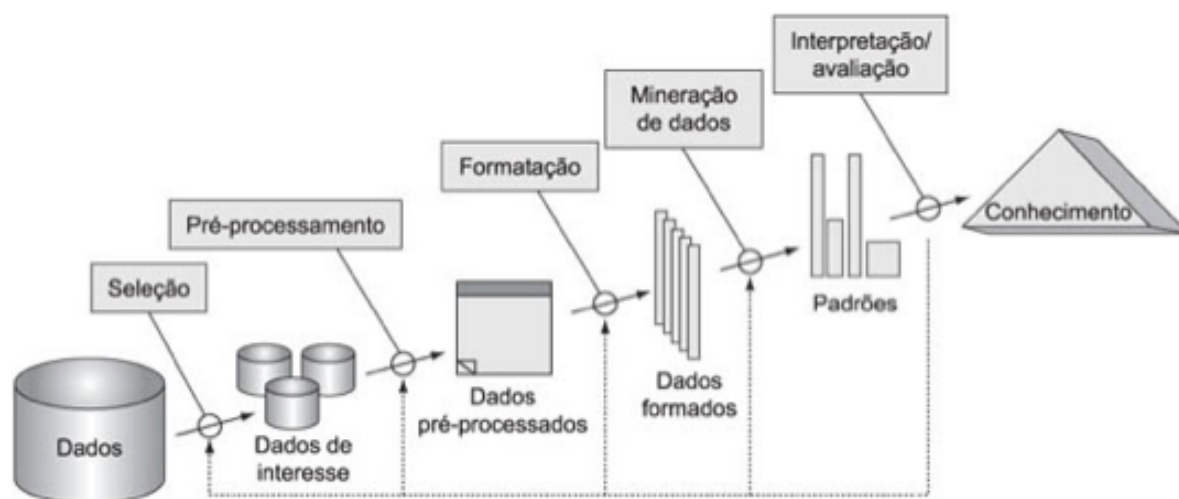
Segundo (TAN; STEINBACH; KUMAR, 2009) a descoberta de conhecimento:

Combina métodos tradicionais de análise de dados com algoritmos sofisticados para processar grandes volumes de dados. Ela também abriu oportunidades interessantes para se explorar e analisar novos tipos de dados e para se analisar tipos antigos de novas maneiras.

Este processo consiste em uma série de etapas de transformação, pré-processamento dos dados, até chegar à fase de pós-processamento dos resultados. Os dados usados neste processo podem estar armazenados em diversos formatos e serem originários de uma fonte de dados central ou até mesmo distribuída (TAN; STEINBACH; KUMAR, 2009).

Na figura 7, estão representados todos os processos envolvidos na Descoberta de Conhecimento.

Figura 7: Etapas do processo de descoberta de conhecimento em base de dados



Fonte: (FAYYAD et al., 1996)

As principais etapas do processo de KDD serão resumidos a seguir (TAN; STEINBACH; KUMAR, 2009), (WITTEN; FRANK, 2005), (FAYYAD et al., 1996):

- **Pré-processamento:** é a transformação dos dados de entrada em um formato apropriado para a mineração. Resumidamente, o pré-processamento abrange as etapas de integração dos dados de múltiplas fontes, a limpeza dos dados para remoção de ruídos e dados duplicados e a seleção de registros e características que serão relevantes nas etapas seguintes;
- **Mineração de dados:** consiste em investigar grandes quantidades de dados com o objetivo de encontrar padrões previamente desconhecidos, úteis e relacionados entre os dados. Um conjunto de fatores motivou o desenvolvimento de técnicas para a mineração de dados, principalmente a necessidade diária de lidar com bancos de dados, com centenas ou milhares de atributos, sendo necessárias técnicas para o tratamento dessa alta dimensionalidade. Atualmente, há diferentes tarefas de mineração de dados, como classificação, regressão, agrupamento e associação;
- **Interpretação/avaliação:** é a etapa que avalia os resultados da mineração de dados, aplicando medições estatísticas para desconsiderar resultados não legítimos. Geralmente, essa etapa é realizada com o apoio de especialistas que avaliam a qualidade e conhecimento obtido a partir dos modelos de mineração de dados.

A mineração de dados é a principal etapa do processo de KDD, sendo separada em dois tipos de tarefas de aprendizado, as tarefas preditivas e descritivas.

2.3.1 Tarefas Preditivas

Nas tarefas preditivas, o objetivo é encontrar funções que a partir de conjuntos de dados possam ser utilizadas para prever o valor ou classe de novos entrantes, com base nos valores de seus atributos. Para isto, os objetos do grupo de treinamento devem ter os mesmos atributos de entrada e saída. Neste tipo de tarefa, é utilizado o aprendizado supervisionado, no qual, se tem um "supervisor externo", que sabe a saída apropriada para cada objeto (FACELI, 2011).

Abaixo serão explicadas as principais tarefas preditivas: classificação e regressão (AMO, 2004), (FACELI, 2011);

- **Classificação:** processo de encontrar um grupo de funções que caracterizam e diferenciam as classes, com a intenção de utilizar o modelo proposto, para prever a classe de objetos que ainda não foram classificados. O modelo concebido, apoia-se

na análise antecipada do conjunto de dados de amostragem ou dados de treinamento, onde estão armazenados os objetos que estão corretamente classificados. Um exemplo de classificação seria determinar se um cliente é bom ou mau pagador.

- **Regressão:** Utiliza-se regressão, quando estamos interessado em descobrir valores que estão faltantes no conjunto de dados, ao contrário de, prever a classe dos objetos. A regressão é indicada quando os valores ausentes são numéricos. Um exemplo de regressão consiste, quando se faz necessário determinar o preço futuro de uma ação na bolsa de valores ou estimar a pressão ideal de um paciente baseado na idade, sexo e massa corporal do mesmo.

2.3.2 Tarefas Descritivas

Nas tarefas descritivas, o intuito é investigar ou detalhar um conjunto de dados. Nestes algoritmos a classe de cada objeto de treinamento é desconhecida, assim como, o número ou conjunto de classes que compõe a base de dados, por esse motivo, se tem uma aprendizagem não-supervisionada (FACELI, 2011), (WITTEN; FRANK, 2005).

Abaixo serão explicadas as principais tarefas descritivas: associação e agrupamento (CAMILO; SILVA, 2009), (AMO, 2004), (TAN; STEINBACH; KUMAR, 2009).

- **Agrupamento:** Ao contrário da classificação e da regressão, onde os objetos que compõe o conjunto de dados já estão classificados, o agrupamento é aplicado sob dados que não possuem suas classes definidas e busca separar elementos com características diferentes em subgrupos ou segmentos com aspectos semelhantes. Um exemplo de uso na análise de grupo é a possibilidade de agrupar clientes que possuam características parecidas formando assim grupos de clientes.
- **Regras de Associação:** O objetivo da análise de associação é descobrir que atributos estão relacionados em uma determinada base de dados. Quando descobertos, os resultados podem ser representados a partir de regras de associação, que são, definidas no formato $X \rightarrow Y$, onde X e Y são elementos do conjunto de dados. Um exemplo de associação encontra-se em uma papelaria, onde um grupo de clientes que compra cadernos também compra canetas, este padrão representa uma regra de associação, no qual um padrão de comportamento repete-se sucessivamente.

2.4 Sabedoria das Massas

Conforme (SUROWIECKI, 2005), sabedoria das massas do inglês *Wisdom of Crowds*, o conhecimento que emerge de decisões coletivas, é em muitos casos mais preciso quando comparado a opinião de um único especialista. No âmbito da descoberta de conhecimento em base de dados, a afirmação anterior, também se faz verdadeira, em virtude da complexidade na escolha de algoritmos, visto que, o melhor algoritmo varia de aplicação para aplicação (DOMINGOS, 2012).

De acordo com (POLIKAR, 2006), (BAZZAN, 2014), esta dificuldade fica mais evidente pelo fato de não termos o conhecimento necessário para selecionar de forma prévia o algoritmo mais apropriado para determinada tarefa. Outros fatores também podem influenciar na escolha, como: quantidade e qualidade dos dados, uma vez que provocam perda de informação.

Buscando a minimização dos problemas citados anteriormente, surge o *Ensemble Learning* ou Conjunto de Máquinas. Este conceito foi primeiramente usado em tarefas de reconhecimento de padrões, como estratégia para reduzir a variabilidade entre conjunto de dados (DIETTERICH, 2002).

O *Ensemble Learning* pode ser entendido como um conjunto de algoritmos que produzem soluções individuais e que de alguma forma combinam suas saídas produzindo uma resposta final. Ainda de acordo com (DIETTERICH, 2002) estudos experimentais mostraram que o uso de um conjunto de algoritmos, em vários casos, é mais preciso quando comparados com métodos que produzem somente uma única resposta.

Por intermédio desse conceito a concepção de extração de informação se transforma, ou seja, ao invés de olhar para o melhor conjunto de dados e o melhor classificador, agora deve-se observar o melhor conjunto de algoritmos, posteriormente, o melhor método de combinação ((DOMINGOS, 2012).

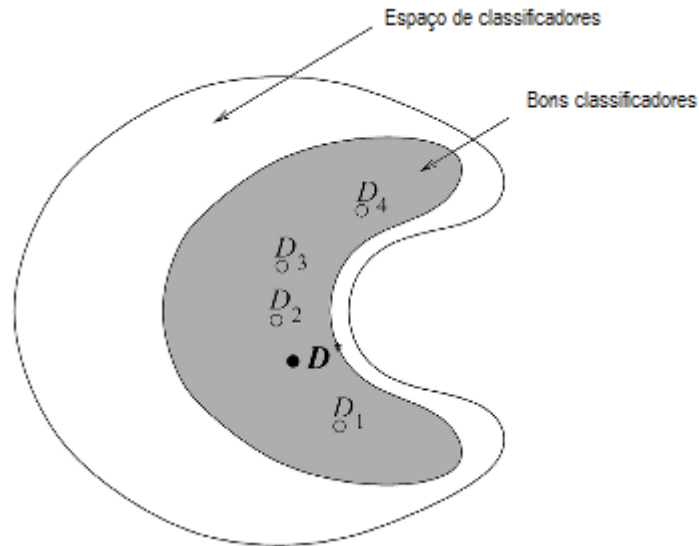
2.4.1 Justificativas para o uso do *Ensemble Learning*

A adoção de sistemas de *Ensemble Learning* em prejuízo de sistemas individuais justifica-se por três razões, sendo elas: estatística, computacional e representabilidade (DIETTERICH, 2002), (BAZZAN, 2014), (KUNCHEVA, 2004), (BARBAT, 2015).

1. **Estatística:** um algoritmo de aprendizado busca encontrar a melhor solução em um determinado espaço de hipóteses. O problema estatístico surge quando escolhe-se somente um único classificador para resolução de determinado problema, dessa maneira, é aumentado o risco estarmos fazendo um má escolha para a situação.

A aplicação de diversos algoritmos simultâneos e independentes tendem a reduzir a possibilidades de uma escolha incorreta.

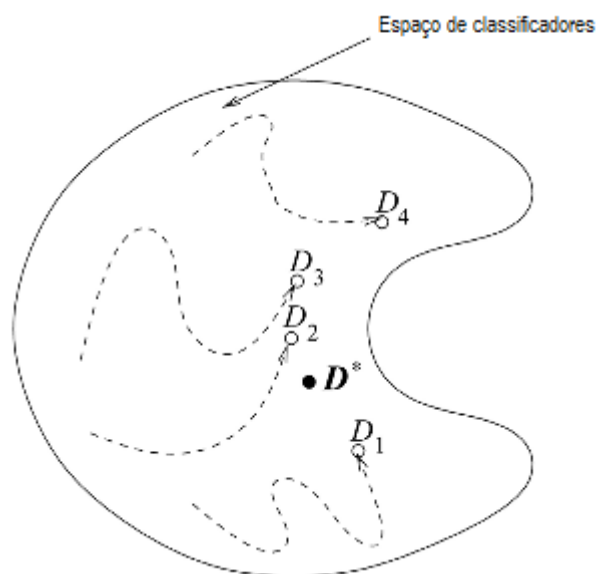
Figura 8: Razão estatística



Fonte: Figura adaptada de (KUNCHEVA, 2004)

- Computacional:** algoritmos de aprendizagem tem por padrão a busca aleatória de soluções, gerando assim, diferentes soluções ótimas. A união de diversas soluções tende a aproximar uma ótima solução global para o problema.

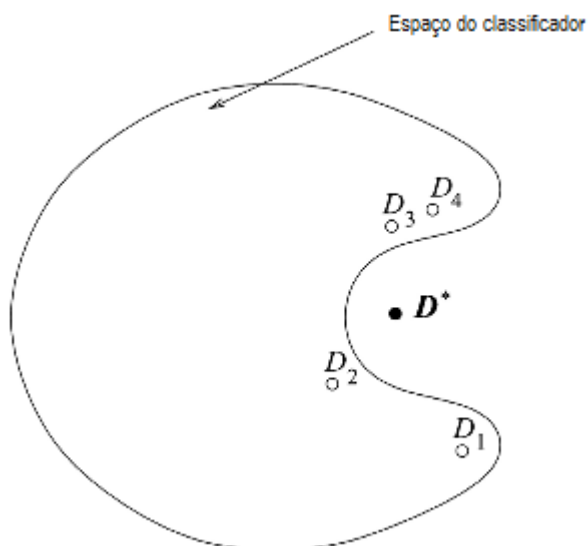
Figura 9: Razão computacional



Fonte: Figura adaptada de (KUNCHEVA, 2004)

3. **Representabilidade:** existe a possibilidade do espaço de hipóteses considerado para o algoritmo de aprendizado não possuir uma solução ótima para o problema, ou seja, os algoritmos de aprendizagem disponíveis não são os adequados para a solução do problema proposto, um exemplo acontece, quando usamos algoritmos não-lineares na solução de problemas lineares.

Figura 10: Razão representacional



Fonte: Figura adaptada de (KUNCHEVA, 2004)

2.4.2 Construção de uma solução com *ensemble learning*

Conforme (KUNCHEVA, 2004) há duas principais formas para produção de sistemas baseados em *Ensemble Learning*. No primeiro caso, cada algoritmo tem a possibilidade de atuar sob todo o espaço de dados, de maneira independente, resultando assim, um conjunto de soluções individuais, com taxas diferentes de precisão. A segunda abordagem consiste na manipulação do conjunto de dados para a geração de múltiplas hipóteses. Através deste comportamento os algoritmos serão executados diversas vezes, utilizando em cada momento uma distribuição diferente de dados. Esta técnica funciona especialmente para algoritmos de aprendizagem instáveis, ou seja, aqueles que a resposta sofre grandes mudanças em resposta a pequenas alterações nos dados de treinamento (FACELI, 2011).

A união das soluções propostas pelas duas abordagens pode ocorrer de diversas maneiras, alguns dos métodos selecionados encontram-se listados abaixo (DIETTERICH, 2002), (BAZZAN, 2014), (KUNCHEVA, 2004), (BARBAT, 2015).

- **Voto da maioria:** provavelmente a estratégia mais intuitiva, neste caso, a decisão é tomada com base na frequência de cada classe, ou seja, a classe que tiver mais votos

será escolhida como decisão final do conjunto de algoritmos. Este processo suporta três padrões de votação, sendo eles: unanimidade, maioria simples e a pluralidade. O primeiro caso acontece quando todos os envolvidos no processo de voto elegem o mesmo candidato, na maioria simples $50\% + 1$ dos votantes escolhem o mesmo postulante, por fim, a pluralidade ocorre quando existe mais de dois candidatos e considera-se vencedor aquele que obtiver a maior quantidade de votos.

- **Voto ponderado da maioria:** é comum que diferentes algoritmos apresentem distintos níveis de acurácia, com isso, aqueles com maior precisão recebem maior peso no processo de votação. O peso de cada algoritmo, é definido de acordo com uma informação prévia ou com base no desempenho dos algoritmos no teste independente.
- **Combinação algébrica:** Neste caso a decisão é tomada com base em qualquer função algébrica. Algumas funções são frequentemente usadas, como: regra do soma, regra da média, regra do produto, regra do máximo, regra do mínimo entre outros.

3 Trabalhos Relacionados

As seções anteriores deste trabalho expuseram fundamentos que auxiliam no entendimento deste trabalho como um todo. No entanto, este capítulo apresenta trabalhos correlatos sobre os temas envolvidos na construção do modelo proposto nesta dissertação.

A seção 3.1 foi desenvolvida com o intuito de obter um panorama referente aos estudos que estão sendo realizados e que usam tecnologias de monitoramento e técnicas de descoberta de conhecimento como objeto de pesquisa. Conjuntamente busca evidenciar o método abordado para a busca de trabalhos relacionados, bem como, os repositórios e palavras-chave escolhidas para a busca de publicações correlacionadas.

As seções 3.2, 3.3 e 3.4 apresentam de forma resumida os trabalhos considerados relevantes para a proposta do presente trabalho.

3.1 Definições e triagens

No processo de pesquisa se faz necessário a definição de termos que reflitam o estudo, bem como, métodos e ambientes de consulta. Esta revisão tem a intenção de obter um conhecimento referente a estudos desenvolvidos que usam avicultura, RFID, descoberta de conhecimento em base de dados e *ensemble learning* como objeto de pesquisa.

Como um dos intuitos desta revisão é verificar as pesquisas na área de uma maneira global, foi escolhido um repositório nacional, juntamente com dois internacionais. Os repositórios estão especificados na tabela 1.

Tabela 1: Repositório selecionados para pesquisa

Repositório	Origem
Alice	Nacional
IEEE	Internacional
CEA	Internacional

Fonte: Autor

O repositório nacional Alice¹ da Embrapa (Empresa Brasileira de Pesquisa Agropecuária) foi escolhido por tratar de um dos assuntos relevantes abordado neste trabalho, a avicultura.

¹ Repositório Alice. Acesso em: <<https://www.alice.cnptia.embrapa.br/>>

Os repositórios internacionais foram selecionados considerando a grande abrangência de anais de eventos e revistas, como também, pela qualidade elevada das publicações, deste modo, foram escolhidos: o repositório Instituto de Engenheiros Eletricistas e Eletrônicos² (IEEE) e *Computers and Eletronics in Agriculture*³ (CEA) da editora Elsevier.

Os termos de busca utilizados na pesquisa estão descritos na tabela 2 considerando suas versões em português (coluna da esquerda) e as respectivas traduções e interpretações em inglês (coluna da direita).

Tabela 2: Palavras-chave escolhidas para o processo de busca

Palavras-chave em português	Palavras-chave em inglês
Avicultura	Aviculture
RFID	RFID
Mineração de dados	Data mining
Ensemble Learning	Ensemble Learning
Frangos de corte	Broiler

Fonte: Autor

Foi definido como intervalo de tempo para seleção das publicações o ano de 2010 até 2015. O resultado inicial dos artigos por meio de uma busca simples, resultou nas quantidades mostradas na Tabela 3.

Tabela 3: Consulta inicial de trabalhos utilizando termos em português

Termos de Busca	Repositórios		
	Alice	IEEE	CEA
Avicultura	236	0	10
RFID	71	9023	5433
Mineração de dados	101	0	19
Ensemble Learning	15	1865	46
Frangos de corte	206	0	4
Total	629	10888	5512

Fonte: Autor

² Repositório do Instituto de Engenheiros Eletricistas e Eletrônicos. Acesso em: <<http://ieeexplore.ieee.org/Xplore/home.jsp>>

³ Revista Computers and Eletronics in Agriculture. Acesso em: <<http://www.journals.elsevier.com/computers-and-electronics-in-agriculture>>

Para a busca com os termos em inglês, o resultado esta representado na tabela 4.

Tabela 4: Consulta inicial de trabalhos utilizando termos em inglês

Termos de Busca	Repositórios		
	Alice	IEEE	CEA
Aviculture	8	0	78
RFID	71	9023	5433
Data Mining	138	29332	21579
Ensemble Learning	15	1865	46
Broiler	279	45	56
Total	519	40265	27192

Fonte: Autor

Após uma primeira análise dos artigos, considerando a ocorrência exata dos termos de pesquisas nos repositórios expostos na tabela 1, bem como, a existência dos mesmos no intervalo de tempo determinado (2010 – 2015), é destacado uma ampla quantidade de publicações a respeito dos assuntos abordados no presente trabalho, bem como, a grande discrepância de resultados entre os idiomas inglês e português.

Com intuito de refinar e obter melhores resultados nas buscas, optou-se por utilizar a combinação de termos, formando assim grupos de palavras-chave para serem pesquisadas nos repositórios. Na primeira pesquisa (tabela 3 e tabela 4) observou-se que para algumas palavras, não foram retornadas publicações que satisfizessem os critérios da pesquisa, desse modo, alguns conjuntos de consultas deixaram de ser executados, visto que, não retornariam trabalhos correlatos. As tabelas seguintes mostram a relação de consultas feitas por repositórios e seus respectivos resultados.

Tabela 5: Consulta a partir da combinação de dois termos em língua portuguesa

Termos de Busca	Repositórios		
	Alice	IEEE	CEA
Avicultura AND RFID	3	0	0
Avicultura AND Mineração de dados	1	0	0
Avicultura AND Ensemble Learning	2	0	0
Avicultura AND Frangos de corte	81	0	0
RFID AND Mineração de dados	5	0	2
RFID AND Ensemble Learning	1	375	65
RFID AND Frangos de Corte	6	0	0
Mineração de dados AND Ensemble Learning	4	7	0
Mineração de dados AND Frangos de corte	2	0	0
Ensemble Learning AND Frangos de corte	2	0	0
Total	107	382	67

Fonte: Autor

Na tabela 5, pode-se observar que o repositório Alice da Embrapa em sua primeira pesquisa, com combinação de termos, retornou foi 107 trabalhos, sendo a maior parte do idioma português.

Tabela 6: Consulta a partir da combinação de dois termos em língua inglesa

Termos de Busca	Repositórios		
	Alice	IEEE	CEA
Aviculture AND RFID	0	0	0
Aviculture AND Data mining	0	0	0
Aviculture AND Ensemble Learning	0	0	0
Aviculture AND Broiler	3	1	10
RFID AND Data Mining	3	132	498
RFID AND Ensemble Learning	1	365	65
RFID AND Broiler	1	7	12
Data mining AND Ensemble Learning	6	6807	2045
Data mining AND Broiler	2	0	0
Ensemble Learning AND Broiler	1	3	6
Total	17	7315	2626

Fonte: Autor

Na tabela 6, é destacado a vasta quantidade de trabalhos, quando comparado aos repositórios nacionais. O repositório **IEEE** apresentou o melhor resultado, retornando 7315 publicações correlatas, destacando-se entre as mesmas, o conjunto de termos **Data Mining and Ensemble Learning** com 6807 trabalhos retornados.

Com base nos resultados das tabelas anteriores, se faz necessário um maior refinamento na busca, visando obter uma quantidade menor de publicações, bem como, trabalhos cada vez mais similares a esta proposta de qualificação.

Tabela 7: Consulta a partir da combinação de três termos em língua portuguesa

Termos de Busca	Repositórios		
	Alice	IEEE	CEA
Avicultura AND RFID AND Mineração de dados	1	0	0
Avicultura AND RFID AND Ensemble Learning	0	0	0
Avicultura AND RFID AND Frangos de corte	2	0	0
Avicultura AND Mineração de dados AND Ensemble Learning	1	0	0
Avicultura AND Mineração de dados AND Frangos de corte	1	0	0
Avicultura AND Ensemble Learning AND Frangos de corte	1	0	0
RFID AND Mineração de dados AND Ensemble Learning	1	0	0
RFID AND Mineração de dados AND Frangos de corte	2	0	0
RFID AND Ensemble Learning AND Frangos de corte	1	0	0
Mineração de dados AND Ensemble Learning AND Frangos de corte	1	0	0
Total	6	0	0

Fonte: Autor

Na tabela 7 estão representados as buscas executadas nos repositórios, bem como, seus respectivos resultados.

Tabela 8: Consulta a partir da combinação de três termos em língua inglesa

Termos de Busca	Repositórios		
	Alice	IEEE	CEA
Aviculture AND RFID AND Data mining	0	0	0
Aviculture AND RFID AND Ensemble Learning	0	0	0
Aviculture AND RFID AND Broiler	0	0	0
Aviculture AND Data mining AND Ensemble Learning	0	0	0
Aviculture AND Data mining AND Broiler	0	0	0
Aviculture AND Ensemble Learning AND Broiler	0	0	0
RFID AND Data mining AND Ensemble Learning	1	140	32
RFID AND Data mining AND Broiler	1	0	4
RFID AND Ensemble Learning AND Broiler	0	2	0
Data mining AND Ensemble Learning AND Broiler	1	1	2
Total	1	143	38

Fonte: Autor

Após a terceira etapa de buscas, representada nas tabelas 7 e 8 é visto uma queda no número de trabalhos relacionados, com relação a tabela 5 e 6, apesar deste resultado, considera-se benéfico para o presente trabalho, uma vez que os trabalhos retornados são cada vez mais relacionados a presente dissertação. Buscando ainda mais, correlação das publicações existentes nos repositórios, com o presente trabalho, é proposto uma nova fase de combinação de termos, procurando aprimorar os resultados.

Tabela 9: Consulta a partir da combinação de quatro termos em língua portuguesa

Termos de Busca	Repositórios		
	Alice	IEEE	CEA
Avicultura AND RFID AND Mineração de dados AND Ensemble Learning	0	0	0
Avicultura AND RFID AND Mineração de dados AND Frangos de corte	1	0	0
Avicultura AND RFID AND Ensemble Learning AND Frangos de corte	0	0	0
Avicultura AND Mineração de dados AND Ensemble Learning AND Frangos de corte	0	0	0
RFID AND Mineração de dados AND Ensemble Learning AND Frangos de corte	0	0	0
Total	1	0	0

Fonte: Autor

Na busca em português, o único repositório que apresentou publicações foi o Alice da Embrapa, porém com uma diminuição de 5 trabalhos, quando comparado com a tabela 7.

Tabela 10: Consulta a partir da combinação de quatro termos em língua inglesa

Termos de Busca	Repositórios		
	Alice	IEEE	CEA
Aviculture AND RFID AND Data mining AND Ensemble Learning	0	0	0
Aviculture AND RFID AND Data mining AND Broiler	0	0	0
Aviculture AND RFID AND Ensemble Learning AND Broiler	0	0	0
Aviculture AND Data mining AND Ensemble Learning AND Broiler	0	0	0
RFID AND Data mining AND Ensemble Learning AND Broiler	0	0	0
Total	1	0	0

Fonte: Autor

A partir da combinação das palavras-chave em conjunto de quatro elementos, percebeu-se uma diminuição substancial na quantidade de trabalhos retornados, tanto na busca em português, quanto em inglês. Por essa razão, decidiu-se não prosseguir com a combinação de termos, visto que, a mesma não retornará trabalhos.

Após a etapa de consulta nos repositórios, por meio do agrupamento de termos. A escolha dos artigos deu-se por meio da análise de seus títulos, resumos e palavras-chave que integram os respectivos trabalhos. Seguindo as definições apresentadas anteriormente, são apresentados na tabela 11 os artigos selecionados para a análise.

Tabela 11: Trabalhos relacionados escolhidos após a etapa de busca

Título	Autores	Universidade	Repositório
Temperature management for the quality assurance of a perishable food supply chain	Myo Min Aung Yoon Seok Chang	Korea Aerospace University	CEA
Wireless sensor networks (WSNs) in the agricultural and food industries	N. Wang Z. Li	Oklahoma State University South China Agricultural University	CEA
Tecnologias da Informação e Comunicação e suas relações com a agricultura	Silvia Maria Fonseca Silveira Massruhá Maria Angelica de Andrade Leite Ariovaldo Luchiar Junior Luciana Alvim Santos Romani	Embrapa Informática Agropecuária	Alice

Fonte: Autor

Da CEA foram selecionados 2 artigos, do Alice 1 artigos. O repositório da IEEE acabou não tendo trabalhos selecionados. Nas seções subsequentes serão apresentados resumos das publicações consideradas relevantes para o presente trabalho.

3.2 *Temperature management for the quality assurance of a perishable food supply chain*

No presente trabalho, os autores (AUNG; CHANG, 2014) apresentam um sistema de monitoramento, capaz de controlar a temperatura de cadeias de abastecimento alimentar. A proposta conforme os autores, surge devido ao limitado tempo de validade que os alimentos perecíveis apresentam, buscando assim, manter a qualidade e reduzir a perdas. O sistema leva em consideração as diferentes temperaturas de armazenamento que os diversos tipos de alimentos necessitam, assim como, os diversificados veículos de transporte.

Para efetivação de controle, os autores propõe um *framework* denominado *Smart Cold Chain Management*, onde cada alimento monitorado recebe uma *tag* RFID, este identificador permitirá ao sistema registrar dados a partir do momento de colheita, armazenamento e transporte do mesmo. Informações acerca da temperatura são colhidas em pequenos intervalos de tempo, não especificados pelo responsável.

Como o principal objetivo da pesquisa era determinar a temperatura ideal para o acondicionamento dos alimentos, os autores definiram como metodologia de escolha, a

mineração de dados. Neste caso, adotou-se a tarefa de agrupamento por meio do algoritmo *K-MEANS*, no qual, buscou encontrar grupos de alimentos que possuíssem temperatura de conservação similares.

Como resultado, os autores alcançaram quatro conjuntos de temperaturas ideais para o armazenamento dos produtos. Os grupos levam em consideração o número de repartições que os veículos de carregamento dispõe. Por fim, os autores relatam que o sistema ajudará, produtores na redução de perdas, assim como na preservação de qualidade do produto.

3.3 *Wireless sensor networks (WSNs) in the agricultural and food industries*

No trabalho proposto por (WANG et al., 2012), os autores traçam um histórico do uso de tecnologias *wireless* na agricultura e indústria alimentícia, expondo seus aspectos positivos e negativos. O primeiro ponto ressaltado pelo autor é o grande número de tecnologias existentes e como sua adoção vêm transformando o ambiente rural, visto que, trata-se de áreas remotas, com grande dependência das condições climáticas.

O autor destaca a versatilidade da tecnologia RFID que pode ser implantada no desenvolvimento de sistemas de rastreamento de alimentos, bem como, no acompanhamento de animais. Ainda de acordo com o autor, o RFID elevou a importância do registro de dados, visto que, existe uma alta demanda por sistemas de rastreabilidade, na qual, não apenas requerem rigorosas inspeções, como também, sistemático recolhimento de informações desde o local de produção até o consumidor. Os pesquisadores salientam que as informações coletadas podem ser usadas para alimentar sistemas de informação, assim como, servirem de entrada para técnicas de análise de dados. Os autores reforçam que a associação de dados históricos e conhecimento agrônomo tendem a melhorar a produtividade, uma vez que com mais informações, maior é o poder de decisão do produtor.

Concluindo o autor, ressalta que o uso de tecnologias no campo tende a crescer, pois demonstram alto potencial de aplicabilidade, do mesmo modo que, apresentam viabilidade financeira.

3.4 Tecnologias da Informação e Comunicação e suas relações com a agricultura

Neste trabalho, os autores (MASSRUBÁ; LEITE et al., 2015) apresentam uma visão geral das possibilidades de inovação através de ferramentas tecnológicas, assim como, suas aplicações na agricultura. Os pesquisadores destacam que as tecnologias da informação vem contribuindo de forma significativa em diversas áreas do conhecimento, permitindo o armazenamento e análise de grandes porções de dados, bem como, possibilitando a troca de informações e conhecimento.

De acordo com responsáveis pelo trabalho, este fenômeno tecnológico também ocorre na agricultura, através do uso de sensores para mapear o solo, monitorar doenças e variáveis climáticas, entre outros casos. O uso dessas aplicações visa obter mais dados sobre a produção, aspectos ambientais e climáticos. Os autores ressaltam que o grande acúmulo de dados tende a auxiliar os produtores, pois com eles, é possível gerar informações e conhecimentos para serem aplicados no setor agrícola.

Por fim, o autor destaca que a adoção de tecnologias no campo tende a aumentar o índice de produtividade, assim como, a eficácia no manejo dos insumos, produzindo assim, redução de custos com mão de obra e aumento na qualidade de trabalho.

4 METODOLOGIA

Neste capítulo são apresentados as ferramentas e métodos utilizados para atingir o objetivo específico do trabalho, ou seja, um modelo classificatório, apto para predição do peso das aves de corte, baseado nas variáveis de temperatura, umidade relativa do ar, luminosidade e idade das aves. Além disso, é apresentado um estudo de caso anterior, com a intenção, de mostrar informações já obtidas sobre o comportamento dos frangos de corte, bem como, apresentar o sistema de informação desenvolvido.

4.1 Tecnologias e linguagens utilizadas no desenvolvimento do trabalho

Neste capítulo são apresentadas as tecnologias e linguagens que foram empregadas no desenvolvimento da proposta.

4.1.1 SQL - *Structured Query Language*

O SQL¹ é a linguagem de programação padrão para gerenciamento e manipulação de bancos de dados relacionais. O SQL usa um conceito de *queries* para inserir, atualizar e modificar dados contidos no banco de dados.

4.1.2 PostgreSQL

O PostgreSQL² é um sistema gerenciador de banco de dados *open-source* que utiliza a linguagem SQL como padrão. Além de ser um software livre, apresenta facilidade no manuseio, bom desempenho e estabilidade.

4.1.3 Arduino

O Arduino³ é uma plataforma *open source* usada para construção de projetos eletrônicos, sendo composta por uma placa física programável (microcontrolador) e uma IDE (*Integrated Development Environment*) utilizada para escrever as instruções que controlaram o microcontrolador.

¹ SQL: http://wiki.icmc.usp.br/images/6/6d/BDAMat08_SQL_DDL.pdf

² PostgreSQL: <https://www.postgresql.org/about/>

³ Arduino: <https://www.arduino.cc/en/Guide/Introduction>

4.1.4 Weka

O Weka⁴ é uma coleção de algoritmos de descoberta de conhecimento em base de dados. Os algoritmos podem ser aplicados diretamente no conjunto de dados. O Weka fornece suporte a todas etapas do processo de extração de informação, como: pré-processamento, mineração de dados (classificação, regressão, associação e agrupamento) e pós-processamento.

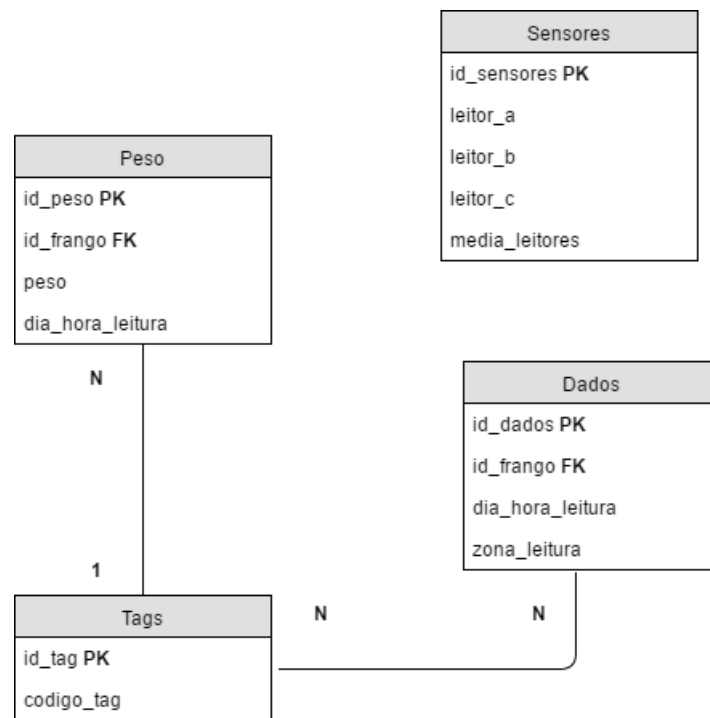
4.2 Modelo entidade-relacionamento (MER)

Nesta seção é apresentada a modelagem do banco de dados que serviu de base para o processo de extração da informação. O objetivo deste banco é propor um relacionamento confiável entre os dados que farão parte do mesmo.

Para a modelagem correta da base de dados, adotou-se o modelo entidade-relacionamento que conforme (PRESSMAN, 2009) baseia-se na percepção do mundo real, sendo constituído de objetos básicos, chamados entidades, e nos relacionamentos entre esses objetos. A proposta do modelo entidade-relacionamento para este trabalho, foi pensado de acordo com as funcionalidades previstas para o sistema.

Na figura 11, encontra-se representado a proposta de MER para o presente sistema.

Figura 11: Modelo entidade-relacionamento



Fonte: Autor

⁴ Weka: <http://www.cs.waikato.ac.nz/ml/weka/>

- **Entidade *Tags*:** armazena dados referentes ao número de identificação do frango no sistema, bem como, o hexadecimal da *tag* que está fixada no frango. Esta tabela relaciona-se com as entidades: **Peso** e **Dados**.
- **Entidade *Sensores*:** foi projetada com a função de armazenar dados a respeito de temperatura, umidade relativa do ar e luminosidade do aviário, assim como, informações de: dia e hora de inserção destes dados na tabela. Esta tabela não relaciona-se com nenhuma outra da base de dados, visto que, sua origem vêm de um arquivo TXT.
- **Entidade *Peso*:** armazena dados referentes ao peso dos animais, bem como, informações do dia da leitura. Esta entidade relaciona-se com a tabela: ***Tags***.
- **Entidade *Dados*:** tem por propósito manter as informações referentes ao monitoramento, buscando armazenar dados sobre o dia e a hora que os frangos encontram-se na zona de leitura da antena RFID. Esta tabela relaciona-se com as entidade: ***Tags***.

4.3 Coleta de dados

Em projetos de pesquisas, o investigador deve relatar de forma detalhada o método de coleta de dados empregado para aquisição das informações. De forma geral, pode-se recolher dados utilizando documentos, observando o comportamento ou até mesmo, por meio de informações fornecidas pelo próprio sujeito de estudo, seja de forma oral (entrevistas) ou por intermédio de questionários (MARCONI; LAKATOS et al., 2002).

A metodologia adotada no presente trabalho apoia-se na observação estruturada em ambiente natural, no qual, o objeto de estudo é observado sem perturbação ou consciência de que esta sendo monitorado. Neste tipo de abordagem são produzidos dados mais confiáveis, visto que, acontece em condições controladas para corresponder a objetivos pré-estabelecidos. Ressalta-se também a posição distante do pesquisador com relação ao objeto analisado, resultando numa interferência quase nula, assim como, a possibilidade do uso de instrumentos no processo de coleta de dados (MARCONI; LAKATOS et al., 2002).

A coleta de dados para o presente trabalho aconteceu na Escola Técnica Estadual Santa Isabel⁵ (ETESI), localizada na zona rural do município de São Lourenço do Sul, Rio Grande do Sul e ocorreu em dois momentos, evidenciados nas seções 4.4 e 4.5. No entanto, é preciso salientar que os equipamentos utilizados para coleta de dados diferem em ambos estudos de caso.

⁵ ETESI: <http://www.etsi.com.br/>

Na tabela 12, encontram-se os aparelhos empregados nas pesquisas.

Tabela 12: Equipamentos utilizados nos estudos de caso

Estudo de Caso 1		Estudo de Caso 2	
Leitor	Motorola XR-450	Leitor	Impinj Speedway Revolution R420
Antena	Polarização Circular	Antena	Polarização Circular
Tags	Ultra High Frequency	Tags	Ultra High Frequency

Fonte: Autor

A adoção de distintos equipamentos, deveu-se a problemas de leitura encontrados na primeira pesquisa, onde o tipo de material da estação de alimentação (metal) apontou uma elevada taxa de interferência nas ondas de radiofrequência emitidas pelo RFID. Desse modo, optou-se pela substituição do leitor para o segundo momento de coleta de dados.

Na figura 12, é possível constatar o local que será disponibilizado para o recolhimento dos dados.

Figura 12: Local disponibilizado para coleta de dados



Fonte: Autor

No presente trabalho, a coleta de dados realizou-se por intermédio da tecnologia RFID explicitada na seção 2.2, enquanto os valores das variáveis de temperatura, umidade relativa do ar e luminosidade foram obtidas, por meio, de sensores incorporados ao Arduino. Os dados para variável de peso foram coletados sempre com intervalo de sete dias, a partir, do início do monitoramento.

No método adotado para coleta de dados, o principal objetivo é verificar o comportamento das aves, com relação a procura por alimentação, desta forma posicionou-se os componentes RFID diretamente nas estações de alimentação dos frangos de corte, desse

modo, sempre que um frango adentrava a zona de leitura da antena, um sinal era enviado para o leitor que, por sua vez, encaminhava os dados obtidos para a base de dados do presente sistema. Todo frango que compôs o estudo recebeu uma *tag*, com intuito de diferencia-lo em relação aos demais (figura 13).

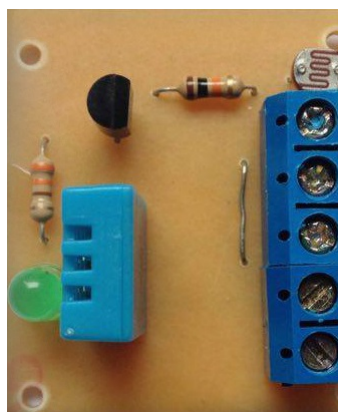
Figura 13: *Tag* RFID no frango



Fonte: Autor

Por fim, os sensores de temperatura e umidade relativa do ar foram dispostos em diferentes pontos do local de estudo, afim de, adquirir dados precisos acerca do tempo. Na figura 14, é possível observar uma das placas posicionados no ambiente.

Figura 14: Sensores usados no desenvolvimento da proposta



Fonte: Autor

A inserção das informações obtidas pelos referidos sensores no banco de dados, ocorrerá sempre que uma mudança nas variáveis climáticas for percebida pelo sistema.

4.4 Estudo de caso 1

O primeiro estudo de caso abordado neste trabalho, busca de forma breve apresentar a análise dos dados coletados num período de 23 dias na Escola Técnica Estadual Santa Isabel, onde objetivou-se compreender a influência que os períodos do dia e o tempo de vida das aves tinham sobre a sua alimentação, por exemplo, uma ave que está no 30º dia de vida, busca se alimentar mais? Em qual período do dia? Neste primeiro experimento as variáveis climáticas não foram abordadas.

Para responder à questão referida acima, foi escolhido o algoritmo de associação Apriori, que segundo Tan, Steinbach e Kumar (2009, p. 17) verifica “[...] se um conjunto de itens é frequente, então todos os seus subconjuntos também devem ser frequentes”. Ainda, conforme os autores, a técnica de análise de dados utiliza a medida de suporte como auxílio para reduzir o número de conjuntos de itens candidatos explorados durante a geração de conjuntos de itens frequentes.

4.4.1 Pré-processamento - Estudo de caso 1

A primeira etapa do processo de descoberta de conhecimento é o pré-processamento de dados, explicado na seção 2.3. No contexto deste trabalho, visando melhorar a qualidade das informações a serem obtidas, por meio, do processo de análise de dados, optou-se por utilizar a discretização, que tem, como objetivo, transformar dados contínuos em dados categóricos. Essa escolha é devido à capacidade de agrupamento de dados contínuos que a discretização oferece.

Dessa forma, a coluna do banco de dados (tabela 13), em que se encontravam as informações referentes aos horários de entrada de cada frango na área de alimentação, foi separada em períodos do dia, sendo eles: madrugada, manhã, tarde e noite.

Tabela 13: Tabela antes do pré-processamento dos dados (Dados Fictícios)

ID_Frango	Zona_Leitura	Dia	Hora
14	T	2015-12-03	06:51:11
16	T	2015-12-03	06:51:12
17	T	2015-12-03	06:51:12
14	T	2015-12-03	06:51:17
15	F	2015-12-03	06:51:17
16	F	2015-12-03	06:51:17
17	F	2015-12-03	06:51:17
14	T	2015-12-03	06:51:22
15	F	2015-12-03	06:51:22
16	T	2015-12-03	06:51:22

Fonte: Autor

Para cada período, foi determinado um intervalo de horário específico, sendo:

- Madrugada: entre 00:00:00 e 05:59:59;
- Manhã: entre 06:00:00 e 11:59:59;
- Tarde: entre 12:00:00 e 17:59:59;
- Noite: entre 18:00:00 e 23:59:59;

A segunda coluna do banco de dados a ser transformada, armazenava informações relacionadas aos dias em que cada frango se encontrava ou não na zona de alimentação. Os atributos desta coluna foram separados em três fases, cada qual correspondendo a um intervalo de tempo. As fases ficaram organizadas da seguinte maneira:

- Fase 1: 8 dias (Segunda semana de vida dos frangos de corte);
- Fase 2: 8 dias (Terceira semana de vida dos frangos de corte);
- Fase 3: 7 dias (Quarta semana de vida dos frangos de corte);

Na tabela 14, é possível observar o novo formato dos dados contidos no banco onde, a coluna que armazenava a identificação de cada frango no sistema foi retirada, visto que, poderia prejudicar o processo de obtenção de informação. Além disso, a coluna que refere-se a presença dos frangos na zona de alimentação foi deslocada para a última posição transformando-se no atributo alvo.

Tabela 14: Tabela após o pré-processamento dos dados (Dados Fictícios)

Fase	Período	Zona de alimentação
Fase 1	Madrugada	T
Fase1	Madrugada	F
Fase1	Madrugada	T
Fase 2	Manhã	F
Fase 2	Manhã	T
Fase 2	Tarde	F
Fase 3	Tarde	T
Fase 3	Noite	F
Fase 3	Noite	T

Fonte: Autor

4.4.2 Mineração de dados - Estudo de caso 1

Após o conjunto de dados ter sido devidamente pré-processado, o mesmo foi disponibilizado para a etapa de mineração de dados esclarecida na seção 2.3. Como explicado anteriormente, o algoritmo escolhido para esta tarefa de mineração de dados foi o *Apriori*, presente no Weka. Para a geração de regras de associação, dois pontos tiveram que ser definidos, sendo eles o suporte e a confiança, sendo o primeiro definido em 0,06 e o segundo em 0,5.

Tanto o uso do suporte quanto o da confiança são importantes, pois garantem que as regras geradas sejam confiáveis. Para explicar melhor o conceito de ambos, Tan, Steinbach e Kumar (2009, p. 17) registram que:

O suporte é uma medida importante porque uma regra que tenha baixo suporte pode acontecer simplesmente por coincidência (...). Por estas razões, o suporte é muitas vezes usado para eliminar regras sem interesse (...). A confiança por outro lado, mede a confiabilidade da inferência feita por uma regra. Para determinada regra $X \Rightarrow Y$, quanto maior a confiança, maior a probabilidade de que Y esteja presente em transações que contenham X.

A partir dessa definição, o algoritmo pode ser executado, gerando as seguintes regras:

1. **Se** Fase = Fase 1 && Período = Tarde **Então** Entered = *True* (Confiança 0.5);
2. **Se** Período = Noite **Então** Entered = *False* (Confiança 0.5);
3. **Se** Fase = Fase 2 && Período = Manhã **Então** Entered = *False* (Confiança 0.5);
4. **Se** Período = Madrugada **Então** Entered = *True* (Confiança 0.5);
5. **Se** Período = Tarde **Então** Entered = *True* (Confiança 0.5);
6. **Se** Fase = Fase 1 && Período = Manhã **Então** Entered = *True* (Confiança 0.5);
7. **Se** Fase = Fase 3 **Então** Entered = *False* (Confiança 0.5);
8. **Se** Fase = Fase 1 **Então** Entered = *True* (Confiança 0.5);
9. **Se** Período = Manhã **Então** Entered = *False* (Confiança 0.5);
10. **Se** Fase = Fase 2 **Então** Entered = *False* (Confiança 0.5);

Com os resultados obtidos, a última tarefa do processo de descoberta de conhecimento pode ser iniciada.

4.4.3 Pós-processamento - Estudo de caso 1

A etapa de pós-processamento visa a avaliar os resultados obtidos com a mineração de dados. Abaixo serão descritas todas as regras geradas pelo algoritmo *Apriori*:

- Regra 1: Esta regra mostra que os frangos que estão na fase 1 de desenvolvimento, durante o período da tarde eles procuram se alimentar mais;
- Regra 2: Esta regra mostra que durante o período da noite as aves, de modo geral, buscam se alimentar menos;
- Regra 3: Na regra 3 os frangos que estão na fase 2 de desenvolvimento, durante o período da manhã buscam se alimentar menos;
- Regra 4: Esta regra indica que durante o período da madrugada os frangos buscam se alimentar mais;
- Regra 5: Esta regra mostra que durante o período da tarde os frangos procuram se alimentar mais;
- Regra 6: A regra 6 indica que os frangos que estão na fase 1 de desenvolvimento, durante o turno da manhã procuram se alimentar mais;
- Regra 7: Esta regra aponta que as aves que estão na fase 3 de desenvolvimento, busca se alimentar menos;
- Regra 8: Esta regra indica que os frangos que estão na fase 1 de desenvolvimento buscam se alimentar mais;
- Regra 9: Esta regra mostra durante o período da manhã os frangos buscam se alimentar menos;
- Regra 10: Nesta regra os frangos que estão na fase 2, procuram se alimentar menos;

De modo geral, foi observado que os frangos buscam se alimentar mais no período da madrugada e da tarde e menos nos períodos da manhã e da noite. Por último, foi detectado que, com o passar do tempo, conforme os frangos cresciam, a procura por alimentação diminuía de forma progressiva.

4.5 Estudo de caso 2

O segundo estudo de caso referido neste trabalho, trata-se dos dados coletados no período de 20 dias na Escola Técnica Estadual Santa Isabel, onde objetivou-se criar um modelo de predição de peso, com base em variáveis de clima, como: temperatura, umidade relativa do ar e luminosidade do ambiente. Outras condições associadas as variáveis, foram: períodos do dia, tempo de vida e peso das aves.

O monitoramento dos frangos de corte aconteceu a partir do 32º dia de vida dos mesmos. Este tempo justifica-se devido as tags adotadas no estudo serem maiores que as aves empregadas no estudo de caso. Na figura 15 é possível observar os animais em seu estágio inicial de vida.

Figura 15: Fonte de calor externa



Fonte: Autor

A escolha das aves da linhagem Cobb⁶, aconteceu por meio de amostragem aleatória, sendo assim, todas foram selecionados de forma eventual, a partir de um conjunto maior (população) de 90 membros. Outro razão a ser considerada é a pouca variabilidade comportamental apresentada pelo animal, tornando assim, a amostra representativa em termos de comportamento.

Por meio do processo de amostragem aleatória foram selecionados 10 elementos, sendo os mesmos colocados em um ambiente separado dos demais. Neste local foram disponibilizados um bebedouro pendular e um comedouro tubular com capacidade de 20 quilos de ração.

⁶ Cobb Vantress: <http://www.cobb-vantress.com/>

Na figura 16, pode-se ver o local concedido para o estudo.

Figura 16: Ambiente de criação dos frangos



Fonte: Autor

No momento de divisão do lote com relação a população, os identificadores RFID foram posicionados na asa das aves. Na figura 17 é possível observar a fixação das *tags*.

Figura 17: Posicionamento *tag* no estudo de caso 2



Fonte: Autor

Posteriormente ao processo de separação e identificação dos frangos de corte, os pesos dos mesmos foram tomados.

Na tabela 15 pode-se verificar os dados coletados.

Tabela 15: Primeira coleta de peso

ID Frango	ID Tag	Peso
1	0003	1920 gramas
2	0015	1770 gramas
3	0010	2020 gramas
4	0004	1520 gramas
5	0011	1370 gramas
6	0006	1520 gramas
7	0007	1520 gramas
8	0019	1020 gramas
9	0028	1820 gramas
10	0009	1620 gramas

Fonte: Autor

A amostra selecionada apresentou peso médio de 1610 gramas. De acordo com (VANTRESS, 2012), um manual de referência sobre nutrição e desempenho de frangos de corte da linhagem Cobb, os mesmos adotados no estudo de caso. O desempenho da amostra encontra-se abaixo do esperado, visto que, para uma criação mista (macho e fêmea) é esperado um peso médio de 1788 gramas. Seguindo a análise, da totalidade de elementos apenas 30% atingiram peso igual ou superior ao sugerido pelo modelo de referência.

Tabela 16: Segunda coleta de peso

ID Frango	ID Tag	Peso
1	0003	2120 gramas
2	0015	2120 gramas
3	0010	2720 gramas
4	0004	2120 gramas
5	0011	1920 gramas
6	0006	2200 gramas
7	0007	2120 gramas
8	0019	1220 gramas
9	0028	2320 gramas
10	0009	1920 gramas

Fonte: Autor

Na segunda semana de monitoramento, novos dados relativos ao peso das aves foram coletados, os mesmos podem ser observados na tabela 16. A partir da análise dos dados, é possível observar um crescimento aproximando de 24% no peso médio da criação. Mesmo com esse aumento, o desempenho dos frangos de corte ainda encontra-se abaixo do recomendado pelo manual de referência (VANTRESS, 2012) que sugere 2447

gramas. Diferentemente da primeira coleta, na qual, três aves apresentaram peso acima do recomendado pelo manual de referência, nesta última somente uma ave apresentou peso igual ou superior. Na terceira semana de coletada de dados, novas informações foram recolhidas. As mesmas podem ser observadas na tabela 17.

Tabela 17: Terceira coleta de peso

ID Frango	ID Tag	Peso
1	0003	2691 gramas
2	0015	2920 gramas
3	0010	3177 gramas
4	0004	2349 gramas
5	0011	2034 gramas
6	0006	2791 gramas
7	0007	2349 gramas
8	0019	1334 gramas
9	0028	3006 gramas
10	0009	2263 gramas

Fonte: Autor

Na tabela acima, a amostra apresentou um peso médio de 2491 gramas. Mesmo com esse progresso, o desempenho quando comparado ao manual de referência, mostrou-se inferior, visto que, o previsto era de 3102 gramas. Igualmente a pesagem anterior, somente um frango de corte apontou a medida esperada. Na quarta semana de pesquisa, os últimos dados referentes ao peso dos frangos de corte foram colhidos, sendo os mesmos evidenciados na tabela 18.

Tabela 18: Quarta coleta de peso

ID Frango	ID Tag	Peso
1	0003	2762 gramas
2	0015	3027 gramas
3	0010	3320 gramas
4	0004	2521 gramas
5	0011	2181 gramas
6	0006	3026 gramas
7	0007	2514 gramas
8	0019	1395 gramas
9	0028	3202 gramas
10	0009	2334 gramas

Fonte: Autor

Na tabela 18, tem-se o peso médio final da amostra antes do abate, sendo o mesmo de 2628 gramas, estando abaixo da medida esperada em cerca de 70%.

4.5.1 Descoberta de Conhecimento em Base de Dados - Estudo de Caso 2

Buscando propor o modelo classificatório, bem como, compreender como as variáveis ambientais agem perante o desenvolvimento dos frangos, foi adotado o algoritmo M5P que produz como saída, árvores de decisão que conforme (FACELI, 2011)

usa a estratégia dividir para conquistar para resolver um problema de decisão. Um problema complexo é dividido em problemas mais simples, aos quais recursivamente é aplicada a mesma estratégia. As soluções dos subproblemas podem ser combinadas, na forma de uma árvore, para produzir uma solução do problema complexo.

Visando proporcionar maior índice de acurácia, também foi empregada a técnica de sabedoria das massas, elucidada na seção 2.4. A opção por algoritmos de árvores de decisão fundamenta-se em testes realizados com o conjunto de dados atual, assim como, o anterior, onde diversas técnicas de classificação foram utilizadas, destacando-se entre as mesmas, as árvores de decisão.

Outro ponto que contribui para adoção do *ensemble learning*, é a maior efetividade desta técnica quando combinada com algoritmos instáveis que são aqueles, onde o modelo preditivo sofre grandes mudanças em resposta a pequenas alterações no conjunto de dados. (FACELI, 2011).

4.5.2 Pré-processamento - Estudo de caso 2

Seguindo a abordagem do experimento anterior, no presente estudo de caso, decidiu-se transformar os dados contínuos em discretos, com objetivo de aglomerar os dados, facilitando assim, a obtenção e interpretação da informação. Na tabela 19 é exibido, os dados antes da etapa de pré-processamento.

Tabela 19: Tabela antes do pré-processamento dos dados

Frango	Dia_Hora_Monitoramento	Temperatura (°C)	Umidade (%)	Peso	Zona de Leitura
1	2016-10-03 06:51:11	30	80	1560	T
3	2016-10-03 06:51:12	30	80	1700	T
7	2016-10-03 06:51:12	30	80	1920	T
1	2016-10-03 06:51:17	30	80	1560	T
5	2016-10-03 06:51:17	30	80	1340	T
3	2016-10-03 06:51:17	30	80	1700	T
8	2016-10-03 06:51:17	30	80	1420	T
1	2016-10-03 06:51:22	30	80	1560	T
9	2016-10-03 06:51:22	30	80	1120	T
3	2016-10-03 06:51:22	30	80	1700	T

Fonte: Autor

Deste modo, dados referentes aos dias de monitoramento e horário das coleta de dados foram discretizados, respeitando assim intervalos pré-definidos.

Abaixo é possível, observar a nova estruturação.

- Semana 4: 7 dias (Quarta semana de vida dos frangos de corte);
- Semana 5: 7 dias (Quinta semana de vida dos frangos de corte);
- Semana 6: 5 dias (Sexta semana de vida dos frangos de corte);

Destaca-se que os dados relativos aos dias de acompanhamento, levam em consideração, a idade das aves no momento da coleta de dados, discernindo assim, do estudo de caso 1. Contudo as informações relacionadas a horário, respeitaram a mesma divisão do experimento anterior.

- Madrugada: entre 00:00:00 e 05:59:59;
- Manhã: entre 06:00:00 e 11:59:59;
- Tarde: entre 12:00:00 e 17:59:59;
- Noite: entre 18:00:00 e 23:59:59;

Além disso, as variáveis climáticas (temperatura e umidade relativa do ar) também foram discretizadas e passaram assim, a respeitar o intervalo definido pelos autores (TINOCO, 1988) e (ABREU; ABREU, 2005) (tab. 20).

Tabela 20: Intervalos de temperatura e umidade - (TINOCO, 1988) e (ABREU; ABREU, 2005)

Idade em Semanas	Temperatura Ambiente (°C)	Umidade do Ar (%)
1	32 a 35	50 - 70
2	29 a 32	50 - 70
3	23 a 26	50 - 70
4	15 a 25	50 - 70
5	15 a 25	50 - 70
6	15 a 25	50 - 70
7	15 a 25	50 - 70

Fonte: (TINOCO, 1988) e (ABREU; ABREU, 2005)

A opção por este intervalo de discretização, se faz, devido a literatura atual não ter unanimidade, a respeito das condições ideais para a criação dos frangos de corte. Um ponto que contribui para esse desacordo é as diferentes condições climáticas que cada local apresenta, dificultando assim, a manutenção de uma escala uniforme.

Por fim, os dados faltantes relativos ao peso das aves foram definidos e preenchidos manualmente, por meio de um cálculo de média, onde, por exemplo o peso da segunda

semana era somado com a da primeira semana e dividido pelo número de dias. Desse modo, foi possível obter uma noção do crescimento diário dos frangos de corte.

Com base nas regras de discretização apresentadas previamente na tabela 21, é possível observar o novo formato dos dados. Assim como, o experimento anterior, a coluna que armazenava a identificação de cada frango no sistema foi retirada, visto que, prejudicaria o processo de obtenção de informação. Além disso, a coluna que refere-se ao peso dos frangos ao longo do período de monitoramento foi deslocada para última posição, transformando-se no atributo alvo.

Tabela 21: Tabela após o pré-processamento dos dados

Fase	Período	Conf. Temperatura	Conf. Umidade	Luminosidade	Peso
Semana 4	Manhã	Temp. em Desconforto	Um. em Conforto	80	1.5
Semana 4	Tarde	Temp. em Desconforto	Um. em Conforto	60	1.57
Semana 4	Noite	Temp. em Desconforto	Um. em Desconforto	0	1.6
Semana 5	Manhã	Temp. em Conforto	Um. em Desconforto	65	1.68
Semana 5	Tarde	Temp. em Conforto	Um. em Conforto	60	1.7
Semana 5	Noite	Temp. em Desconforto	Um. em Conforto	0	1.71
Semana 5	Manhã	Temp. em Conforto	Um. em Desconforto	70	1.8
Semana 5	Tarde	Temp. em Desconforto	Um. em Conforto	50	1.9
Semana 6	Noite	Temp. em Desconforto	Um. em Desconforto	0	2
Semana 6	Madrugada	Temp. em Conforto	Um. em Desconforto	10	2.04
Semana 6	Manhã	Temp. em Conforto	Um. em Desconforto	50	2.1

Fonte: Autor

4.5.3 Mineração de dados - Estudo de caso 2

Após a etapa de pré-processamento de dados, foi iniciada a fase de mineração de dados, elucidada na seção 2.3. Neste estudo de caso foram empregadas duas abordagens para obtenção da informação, sendo a primeira somente com o algoritmo *MP5*, que constrói modelos baseados em árvores de regressão, onde os nós folhas contém funções lineares (QUINLAN et al., 1992).

No segundo método, decidiu-se utilizar o algoritmo anterior em conjunto, com a técnica de *ensemble learning Bagging*, na qual, cria replicações do conjunto de dados, por meio de amostragem por reposição. Cada amostra retirada do conjunto de dados possui o mesmo tamanho. Após esse processo, cada porção colhida do grupo de dados é treinada e responsável pela geração de um classificador. Todo o classificador obtido é usado para classificar os conjuntos de teste, ao final deste procedimento, é adotado um esquema de voto uniforme para decidir o melhor modelo de classificação. A escolha desta técnica de mineração de dados esta fundamentada em seu desempenho em algoritmos de árvore de decisão, expostas no artigo intitulado "*Bagging Predictors*" de Leo Breiman (BREIMAN, 1996), (FACELI, 2011), (TAN; STEINBACH; KUMAR, 2009).

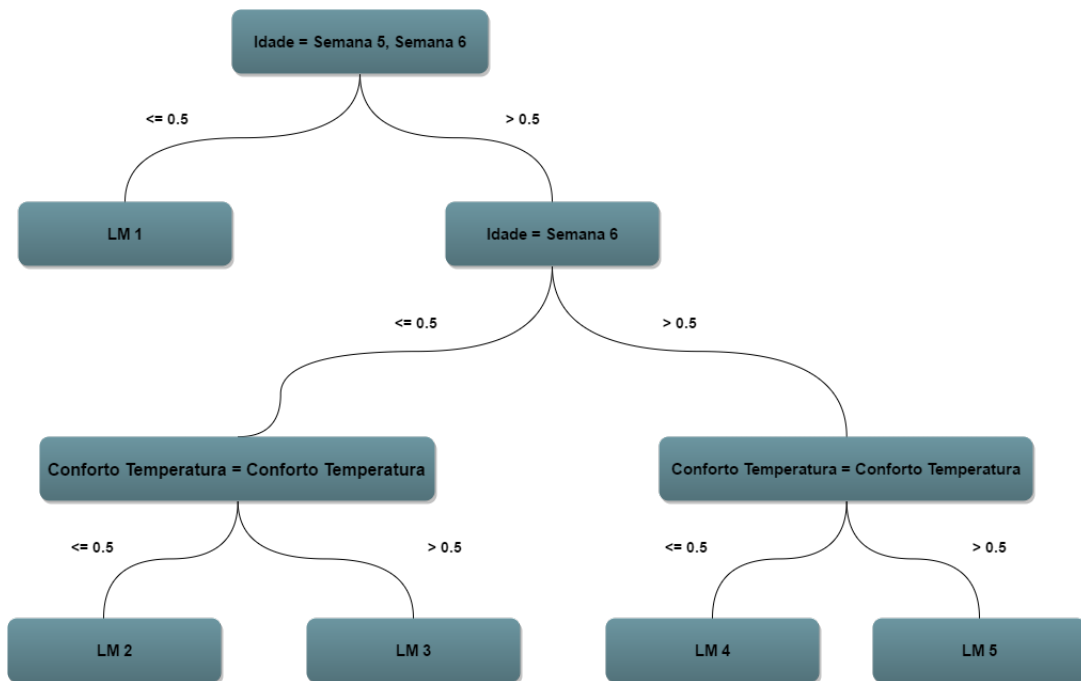
Em ambos experimentos, os algoritmos foram executados com os parâmetros padrões do *software* de mineração de dados Weka, com base nesta escolha, pode-se iniciar

o processo de mineração de dados e conseqüentemente a etapa de pós-processamento.

4.5.4 Pós-processamento - Estudo de caso 2

Na figura 18, é possível observar o modelo gerado pelo algoritmo M5P, como dito previamente, cada nó folha da árvore representa uma função matemática de grau um.

Figura 18: Resultado do algoritmo M5P



Fonte: Autor

Prosseguindo a análise da árvore, pode-se constatar que o atributo de maior importância para obtenção do modelo é o referente ao tempo de vida das aves, visto que, encontra-se na raiz da árvore. O segundo atributo em importância é relativo ao conforto em termos de temperatura dos animais. As propriedades restantes do banco de dados foram desconsideradas para obtenção do modelo, mas não para elaboração das fórmulas, como pode ser constatado, a seguir.

- LM 1:** Média Peso = $0.2659 * \text{Idade} = \text{Semana 5, Semana 6}$
 $+ 0.1041 * \text{Idade} = \text{Semana 6}$
 $+ 35.4978 * \text{Período} = \text{Tarde, Noite, Madrugada}$
 $- 54.2532 * \text{Período} = \text{Noite, Madrugada}$
 $+ 45.6545 * \text{Período} = \text{Madrugada}$
 $+ 0.0022 * \text{Conforto Temperatura} = \text{Temperatura de Conforto}$
 $+ 1897.1283$

- **LM 2:** Média Peso = $0.1817 * \text{Idade} = \text{Semana 5, Semana 6}$
 $+ 0.1869 * \text{Idade} = \text{Semana 6}$
 $- 29.0627 * \text{Período} = \text{Tarde, Noite, Madrugada}$
 $- 16.804 * \text{Período} = \text{Noite, Madrugada}$
 $+ 39.3066 * \text{Período} = \text{Madrugada}$
 $+ 0.0902 * \text{Conforto Temperatura} = \text{Temperatura de Conforto}$
 $+ 2389.9214$
- **LM 3:** Média Peso = $0.1817 * \text{Idade} = \text{Semana 5, Semana 6}$
 $+ 0.1869 * \text{Idade} = \text{Semana 6}$
 $- 48.2634 * \text{Período} = \text{Tarde, Noite, Madrugada}$
 $+ 42.6607 * \text{Período} = \text{Noite, Madrugada}$
 $+ 18.7775 * \text{Período} = \text{Madrugada}$
 $+ 0.2521 * \text{Conforto Temperatura} = \text{Temperatura de Conforto}$
 $+ 2456.5986$
- **LM 4:** Média Peso = $0.1817 * \text{Idade} = \text{Semana 5, Semana 6}$
 $+ 0.254 * \text{Idade} = \text{Semana 6}$
 $- 17.4438 * \text{Período} = \text{Tarde, Noite, Madrugada}$
 $+ 7.4946 * \text{Período} = \text{Noite, Madrugada}$
 $+ 0.0467 * \text{Período} = \text{Madrugada}$
 $+ 0.0317 * \text{Conforto Temperatura} = \text{Temperatura de Conforto}$
 $+ 2584.7963$
- **LM 5:** Média Peso = $0.1817 * \text{Idade} = \text{Semana 5, Semana 6}$
 $+ 0.254 * \text{Idade} = \text{Semana 6}$
 $+ 15.9923 * \text{Período} = \text{Tarde, Noite, Madrugada}$
 $+ 0.0249 * \text{Período} = \text{Noite, Madrugada}$
 $+ 10.8324 * \text{Período} = \text{Madrugada}$
 $+ 0.0048 * \text{Conforto Temperatura} = \text{Temperatura de Conforto}$
 $+ 2559.929$

O entendimento do modelo gerado, se faz, por meio do percorrimento da árvore. Abaixo, é exposto um algoritmo para interpretar a árvore. Uma observação, se faz, importante os valores condicionais ≤ 0.5 e > 0.5 , devem ser lidos como: *false* e *true*, respectivamente.

1. **Primeira regra:** Idade é igual a semana 5 ou semana 6?

- a) Se não, deslocar para esquerda e usar a fórmula linear LM 1.
- b) Se sim, deslocar para direita. Ir para segunda regra.

2. **Segunda regra:** Idade é igual a semana 6?
 - a) Se não, deslocar para esquerda e ir para terceira regra.
 - b) Se sim, deslocar para direita e ir para quarta regra.
3. **Regra Três:** O ambiente esta em temperatura de conforto?
 - a) Se não, deslocar para esquerda e usar a fórmula linear LM 2.
 - b) Se sim, deslocar para direita e usar a fórmula linear LM 3.
4. **Regra Quatro:** O ambiente esta em temperatura de conforto?
 - a) Se não, deslocar para esquerda e usar a fórmula linear LM 4.
 - b) Se sim, deslocar para direita e usar a fórmula linear LM 5.

Supondo que a função a ser executada, seja a LM 4. O peso esperado para um frango de corte, que se encontra na semana 6 de monitoramento, no período da tarde, turno noite e fora da temperatura de conforto é de 2575.2828 gramas, conforme o modelo concebido pelo algoritmo. Comparando o valor real *versus*, o valor esperado, o erro é de 28.9828 gramas.

Prosseguindo a análise do modelo, abaixo tem-se as métricas de avaliação disponibilizadas pelo mesmo. Sendo elas:

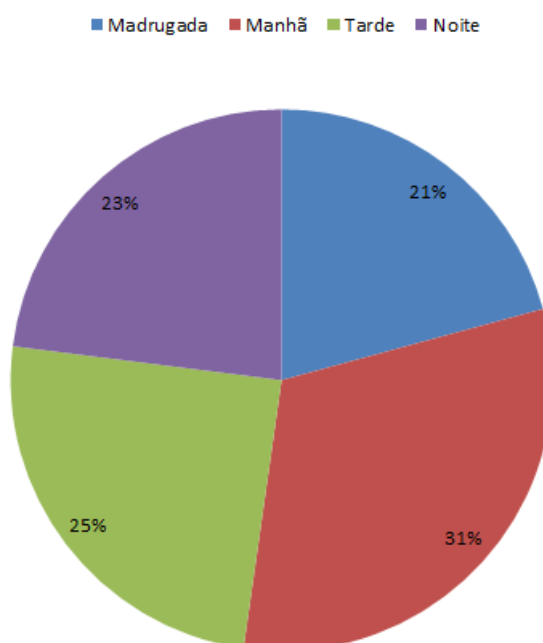
- **Coefficiente de Correlação:** mede o grau de relação linear entre variáveis quantitativas. Este coeficiente oscila entre os valores -1 e 1. O valor 0 expressa há inexistência de relação linear, o valor 1 representa uma correlação linear perfeita, bem como, a medida -1, no entanto, simboliza de forma inversa, ou seja, quando uma variável cresce a outra decresce.
- **Erro Quadrático Relativo:** mede a taxa de erro absoluto do modelo proporcionado pelo algoritmo. Compara o real valor armazenado na base de dados, contra o valor de saída fornecido pelo algoritmo. Quanto mais próximo de 0, melhor é o modelo gerado.

Na primeira execução, onde somente foi utilizada o algoritmo M5P, os valores do coeficiente de correlação e erro quadrático relativo, foram 0,9638 e 26.6642%, respectivamente. Tanto a primeira, quanto a segunda medida de avaliação, foram satisfatórias, visto que, no primeiro caso os dados apresentam um alto grau de correlação, ou seja, os atributos possuem uma forte associação linear.

Como afirmado antecipadamente, o processo de mineração de dados envolveu dois métodos distintos. Na segunda abordagem, o uso da técnica de *ensemble learning*, *Bagging*, também obteve resultados satisfatórios, no entanto, o ganho de informação foi considerado ínfimo, quando comparado com o algoritmo M5P. Na primeira métrica de avaliação considerada pelo presente trabalho, o coeficiente de correlação apresentou uma taxa de 0.9638, igual ao experimento anterior. O ganho diminuto de informação aconteceu na segunda métrica, onde o modelo apresentou o valor de 26.6635%.

Mesmo não fazendo parte do escopo do segundo estudo de caso, algumas informações acerca dos dados também puderam ser obtidas. Como o período de maior frequência das aves na zona de alimentação, exposto na figura 19.

Figura 19: Gráfico de procura por alimentação por períodos do dia

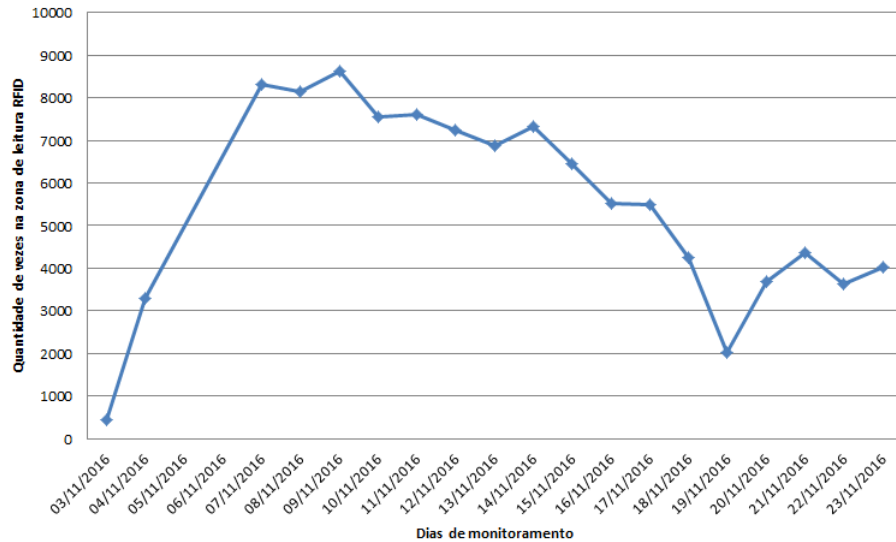


Fonte: Autor

Na figura acima, é possível observar que o período de maior movimentação dos frangos de corte em torno da zona de alimentação, é no turno da manhã, representando 31%. Sendo sucedido pelos horários da tarde, madrugada e noite, respectivamente.

No segundo gráfico (20) exposto no trabalho, temos a representação da evolução alimentar dos frangos de corte em termos de busca por alimentação.

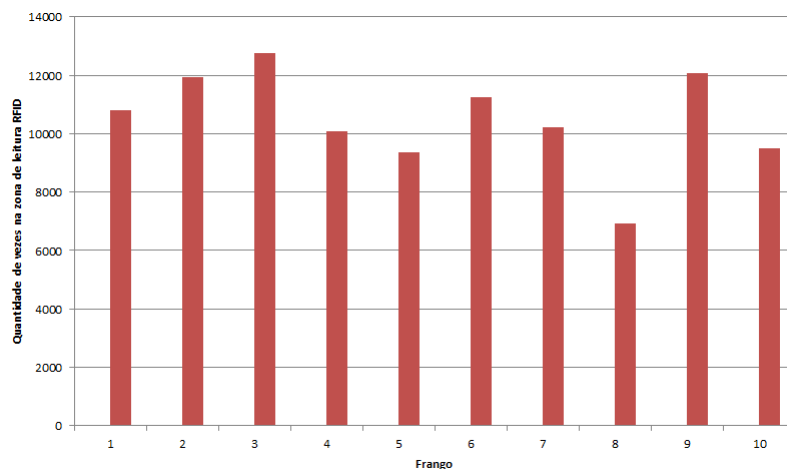
Figura 20: Gráfico de evolução com relação a procura por alimentação



Fonte: Autor

No esquema acima, temos a possibilidade de observar que ao longo do tempo de vida das aves, a busca por alimentação diminui de forma gradativa. Esta reação, é natural, visto que, com o passar do tempo as aves ganham mais peso, adquirindo assim, maiores complicações na locomoção, dificultando o processo de alimentação. No gráfico 21 é mostrado um comparativo entre os frangos de corte em função da variável, busca por alimentação.

Figura 21: Gráfico comparativo dos frangos de corte com relação a busca por alimentação



Fonte: Autor

No gráfico, é possível perceber que com exceção do frango 8, o restante da amostra teve um comportamento regular em termos de frequência na zona de alimento, destacando-

se entre as aves 2, 3 e 9. Esse desempenho, ou seja, uma maior quantidade de vezes na zona de leitura, se reflete em termos de conversão alimentar.

Na tabela 22, pode-se observar a afirmação feita previamente.

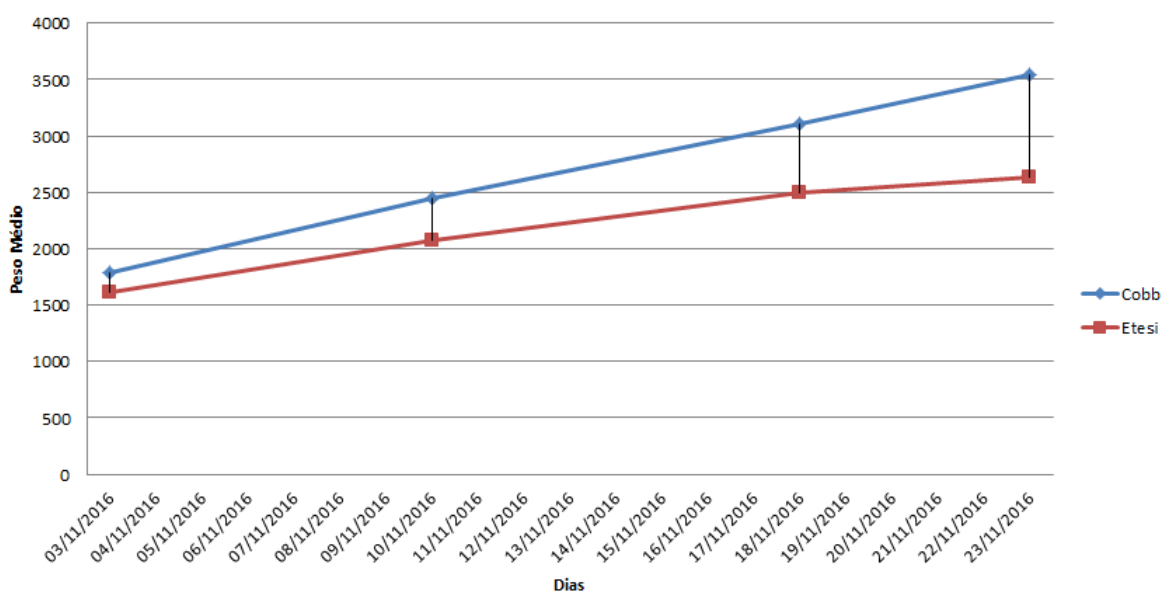
Tabela 22: Conversão alimentar

Frango	Qtd. entradas na zona de alimentação	Peso médio (G)
1	10818	2762
2	11943	3027
3	12766	3320
4	10090	2521
5	9349	2181
6	11235	3026
7	10202	2514
8	6921	1395
9	12067	3202
10	9508	2334

Fonte: Autor

Na tabela acima, fica perceptível que os frangos de corte que passaram maior parte do tempo no interior da zona de leitura RFID tiveram um ganho de peso superior, salvo algumas exceções, como por exemplo: as aves 4 e 7, apresentaram 2521 g e 2514 g, bem como, uma quantidade de 10090 e 10202 de vezes na zona de alimentação, respectivamente. No gráfico 22, temos uma comparação entre a média de peso dos frangos de corte da Etesi *versus* o valor esperado pelo produtor da linhagem.

Figura 22: Ambiente de criação dos frangos

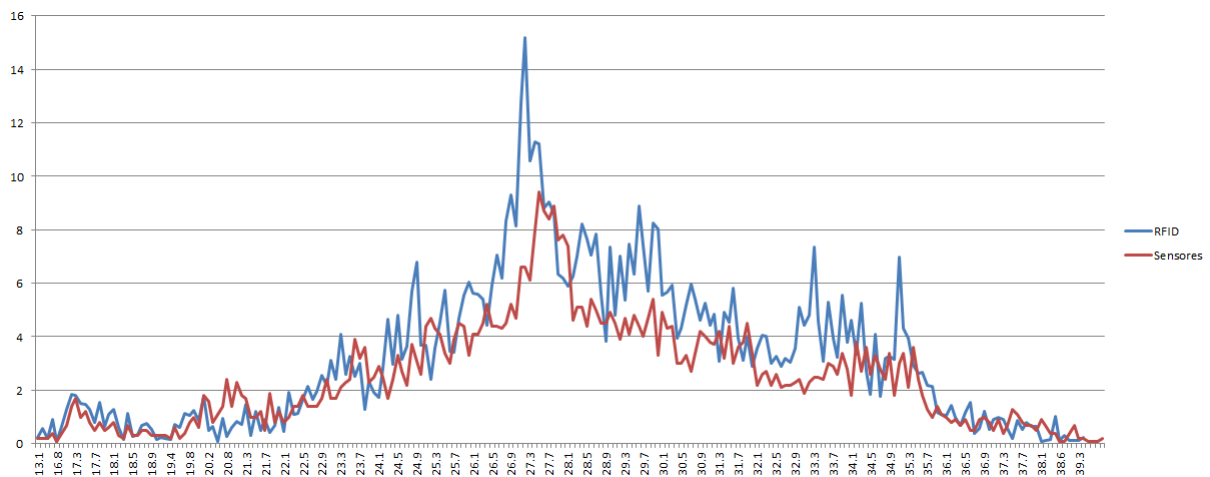


Fonte: Autor

Com base no gráfico, algumas hipóteses podem ser levantadas, afim de, tentar justificar o desempenho abaixo do esperado, como: alimentação deficitária em termos de nutrientes (farelo de milho e concentrado), assim como, temperatura e umidade relativa do ar incompatíveis com as necessidades fisiológicas do animal, visto que, essas variáveis impactam diretamente no desempenho das aves. Influenciando, em termos de busca por alimentação ou conversão alimentar.

Prosseguindo com as análises, no gráfico 23 é mostrada a relação de procura por alimentação dos frangos de corte com as temperaturas registradas pelos sensores. Na linha azul, é exibida a quantidade de vezes que as aves de corte adentraram a zona de leitura em determinada temperatura, enquanto, em vermelho é exposto a quantidade de vezes que cada leitura temperatura foi lida.

Figura 23: Leitura RFID x Sensores Temperatura

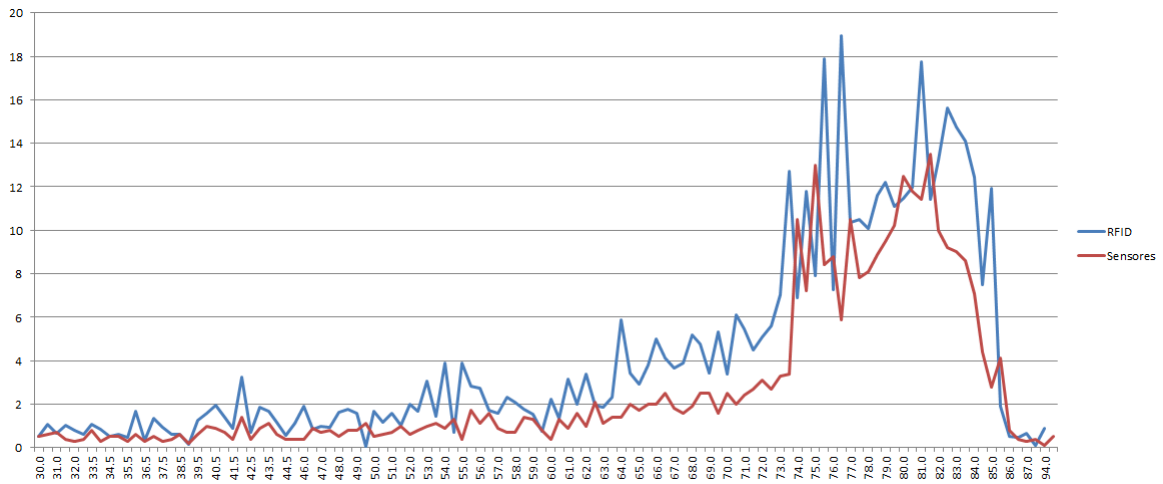


Fonte: Autor

No esquema acima, pode-se observar que a temperatura mais frequente na leitura dos sensores (27°C), trata-se da preferida das aves de corte para procura por alimentação.

No gráfico seguinte, é mostrada a associação da variável umidade com o total de entradas das aves na zona de cobertura RFID, assim como, a imagem anterior, a linha azul expõe o número de vezes que frangos estiveram na zona de leitura em determinada umidade, enquanto, em vermelho é evidenciado os registros de umidade, por parte dos sensores.

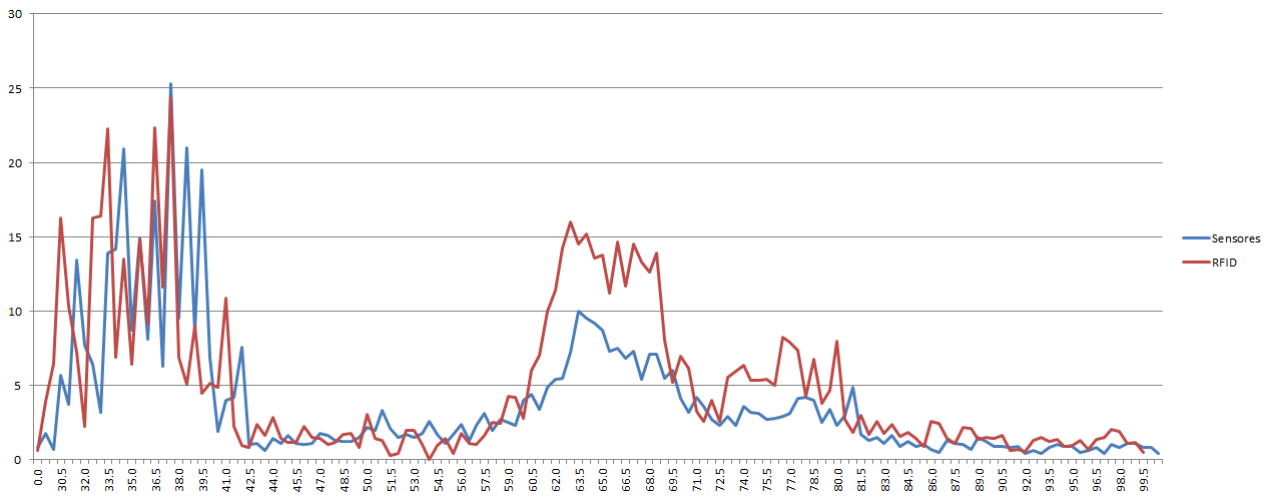
Figura 24: Leitura RFID x Sensores Umidade



Fonte: Autor

Diferentemente do gráfico anterior, o valor de umidade mais constante por parte dos sensores, não representa a preferência das aves em termos de procura por alimentação, ou seja, os frangos de corte buscam alimentar-se num intervalo de umidade distinto (77% - 78%). Por fim, no gráfico 25 é expressado a correlação da variável de luminosidade com a totalidade de entradas dos frangos na área de abrangência do sistema de monitoramento.

Figura 25: Leitura RFID x Sensores Luminosidade



Fonte: Autor

Semelhante ao gráfico 23, o intervalo de luminosidade com maior frequência, entre 33% e 39.5%, refere-se ao momento de maior procura por alimentação dos frangos de corte. Por meio, da análise dos gráficos 23, 24 e 25, é possível reconhecer que as condições ideais para criação dos frangos de corte na Escola Estadual Santa Isabel no período de 03/11/2016 a 23/11/2016, seriam:

- **Temperatura:** entre 26.5°C e 27.3°C;
- **Umidade:** entre 75% e 77%;
- **Luminosidade:** entre 33% e 39.5%;

As condições evidenciadas acima, estão ligadas a análise dos registros do segundo estudo de caso, no entanto, salientasse que os valores de intervalo para as variáveis, pouco diferem do limite superior recomendado pelo manual de referência da linhagem.

5 CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

De forma geral, foi possível observar que o uso da tecnologia de monitoramento em ambos estudos de caso, não gerou interferência no desenvolvimento das aves, como também contribui para aquisição de informações úteis acerca dos hábitos alimentares dos frangos de corte, como: o período de maior movimentação em torno da área de alimentação, dados sobre a evolução alimentar das aves e comparações entre os frangos de corte da amostra. Deve-se destacar que os dois estudos de caso alcançaram os mesmos resultados, no que toca, as informações listadas anteriormente.

Do mesmo modo, a descoberta de conhecimento em base de dados mostrou-se válida, pois através dela, foi possível alcançar o objetivo geral do trabalho, ou seja, a determinação de um modelo de classificação para predição do peso dos frangos de corte, perante variáveis, como: temperatura, umidade relativa do ar, luminosidade, peso e idade das aves. Por meio do modelo foi possível compreender como as variáveis citadas previamente influenciam no processo de desenvolvimento dos frangos de corte, mostrando que um ambiente de criação deficitário em termos de conforto térmico, prejudica o desenvolvimento das aves em termos de conversão alimentar.

Da mesma maneira, é ressaltado o uso do *ensemble learning* no segundo estudo de caso, que apesar de não oferecer ganhos significativos de informação, quando comparado ao uso de um único algoritmo, demonstrou ser uma alternativa válida para o processo de mineração de dados.

Destaca-se que apesar dos resultados satisfatórios, é requerida uma quantidade maior de validações, para provar que os dados obtidos neste trabalho evidenciam a realidade no comportamento das aves de corte e não um padrão fora do normal.

Por fim, salienta-se o caráter multidisciplinar do presente trabalho, buscando não apenas, contribuir na área da computação, mas também em outras áreas como pecuária e agrocomputação. Com propósito de servir de recurso em iniciativas que necessitem/demandem novas tecnologias para o auxílio em tarefas de pesquisa, monitoramento e exploração de ambientes avícolas.

As próximas etapas desse trabalho estão relacionadas a implementação de novas funcionalidades para o processo de coleta de dados. Entre os novos entendimentos, estão:

- Teste de novos algoritmos de mineração de dados;
- Colocação de uma célula de carga diretamente no comedouro, desse modo, será

possível ter o controle do consumo de ração;

- Posicionamento de sensores de umidade do solo, a fim de, verificar a porcentagem de água existente na cama das aves;
- Instalação de sensores de vibração diretamente nas estações de alimentação das aves. Desse modo, além de termos conhecimentos de quais frangos estão na zona de alimentação, teremos o controle de quantos de fato estão se alimentando;
- Aquisição de sensores de velocidade do vento. Dessa maneira será possível ter um controle maior, acerca das condições das aves, em termos de conforto térmico;
- Propor o uso de atuadores no ambiente de coleta de dados, com objetivo, de intervir nas condições do local, trazendo assim um ambiente agradável para os frangos de corte;
- Aperfeiçoamento da aplicação móvel para formulação de ração dos frangos de corte;

Referências

- ABREU, P.; ABREU, V. Diagnóstico bioclimático para produção de aves no centro norte da bahia. In: *XXXIV Congresso Brasileiro de Engenharia Agrícola. Canoas-RS*. [S.l.: s.n.], 2005.
- AGRICULTURA, M. da. *Relatório Anual*. 2015. Disponível em: <<https://goo.gl/ypAb9b>>. Acessado em: 25 de maio de 2016.
- AMO, S. de. Técnicas de mineração de dados. *Jornada de Atualização em Informática*, 2004.
- AUNG, M. M.; CHANG, Y. S. Temperature management for the quality assurance of a perishable food supply chain. *Food Control*, Elsevier, v. 40, p. 198–207, 2014.
- AVICULTURA, U. B. d. *Relatório Anual*. 2016. Disponível em: <<http://goo.gl/4Mg09a>>. Acessado em: 25 de maio de 2016.
- BARBAT, M. *Classificação de sedimentos superficiais de leitos oceânicos e de águas continentais: Uma abordagem baseada em super segmentação e sabedoria das massas*. Tese (Doutorado) — UNIVERSIDADE FEDERAL DO RIO GRANDE, 2015.
- BARWALDT, R. et al. Sustentabilidade tecnológica em cadeia produtiva avícola de corte, através da tecnologia rfid. 2014.
- BAZZAN, A. L. *Exploring ensemble learning techniques to optimize the reverse engineering of gene regulatory networks*. Tese (Doutorado) — UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL, 2014.
- BREIMAN, L. Bagging predictors. *Machine learning*, Springer, v. 24, n. 2, p. 123–140, 1996.
- CAMILO, C. O.; SILVA, J. C. d. Mineração de dados: Conceitos, tarefas, métodos e ferramentas. *Universidade Federal de Goiás (UFG)*, p. 1–29, 2009.
- DIETTERICH, T. G. Ensemble learning. *The handbook of brain theory and neural networks*, MIT Press: Cambridge, MA, v. 2, p. 110–125, 2002.
- DOMDOUZIS, K.; KUMAR, B.; ANUMBA, C. Radio-frequency identification (rfid) applications: A brief introduction. *Advanced Engineering Informatics*, Elsevier, v. 21, n. 4, p. 350–355, 2007.
- DOMINGOS, P. A few useful things to know about machine learning. *Communications of the ACM*, ACM, v. 55, n. 10, p. 78–87, 2012.
- FACELI, K. *Inteligência Artificial: Uma abordagem de aprendizado de máquina*. [S.l.]: Grupo Gen-LTC, 2011.
- FAYYAD, U. M. et al. *Advances in knowledge discovery and data mining*. the MIT Press, 1996.

- GIROTTO, A.; AVILA, V. Sistema de produção de frangos de corte. *EMBRAPA Suínos e Aves. Versão Eletrônica*, 2003.
- GOMES, H. M. C. Construção de um sistema de rfid com fins de localização especiais. Universidade de Aveiro, 2007.
- KUNCHEVA, L. I. *Combining pattern classifiers: methods and algorithms*. [S.l.]: John Wiley & Sons, 2004.
- MARCONI, M. d. A.; LAKATOS, E. M. et al. *Técnicas de pesquisa*. [S.l.]: São Paulo: Atlas, 2002.
- MASSRUBÁ, S. M. F. S.; LEITE, M. A. d. A. et al. *Tecnologias da Informação e Comunicação e suas relações com a agricultura*. [S.l.]: Brasília: Empbrapa, 2015.
- MENDES, A. A. *Panorama da avicultura nacional e perspectivas do setor*. 2014. Disponível em: <<https://goo.gl/uX113A>>. Acessado em: 12 de janeiro de 2017.
- POLIKAR, R. Ensemble based systems in decision making. *Circuits and systems magazine, IEEE, IEEE*, v. 6, n. 3, p. 21–45, 2006.
- PRESSMAN, R. S. *Engenharia de software*. [S.l.]: AMGH Editora, 2009.
- QUINLAN, J. R. et al. Learning with continuous classes. In: SINGAPORE. *5th Australian joint conference on artificial intelligence*. [S.l.], 1992. v. 92, p. 343–348.
- RODRIGUES, A. S. d. P. Inserção de tecnologia no processo de monitoramento em cadeia produtiva avícola. 2014.
- SILVA, M. P. d. S. Mineração de dados: Conceitos, aplicações e experimentos com weka. *Sociedade Brasileira de Computação*, v. 1, 2004.
- SORJ, B.; POMPERMAYER, M. J.; CORADINI, O. L. Camponeses e agroindústria: transformação social e representação política na avicultura brasileira. Centro Edelstein, 2008.
- SUROWIECKI, J. *The wisdom of crowds*. [S.l.]: Anchor, 2005.
- TAN, P.-N.; STEINBACH, M.; KUMAR, V. *Introdução ao datamining: mineração de dados*. [S.l.]: Ciencia Moderna, 2009.
- TINOCO, I. d. F. F. *Resfriamento adiabático (evaporativo) na produção de frangos de corte*. Tese (Doutorado) — Universidade Federal de Viçosa., 1988.
- VANTRESS, C. *Suplemento: Desempenho e Nutrição para Frangos de Corte*. 2012. Disponível em: <<https://goo.gl/I9U4Bz>>. Acessado em: 28 de dezembro de 2016.
- WANG, N. et al. Wireless sensor networks (wsns) in the agricultural and food industries. *Robotics and Automation in the Food Industry: Current and Future Technologies*, Elsevier, p. 171, 2012.
- WITTEN, I. H.; FRANK, E. *Data Mining: Practical machine learning tools and techniques*. [S.l.]: Morgan Kaufmann, 2005.

Apêndices

APÊNDICE A – Sistema de Informação

O funcionamento do sistema desenvolvido para gerenciamento de aviários, pode ser exemplificado, por meio de duas etapas, na qual, a primeira possibilita o usuário coletar informações a respeito do aviário, através do RFID. A partir deste ponto, sem a necessidade da coleta de dados, algumas funcionalidades são implementadas, como: cadastro/exclusão de usuários, aviários e frangos, assim como, a utilização do sistema para visualizar relatórios. Para melhor entendimento das funcionalidades implementadas no sistema, abaixo encontra-se a descrição das mesmas.

- **Cadastro de usuários:** nesta opção o usuário do sistema tem a possibilidade de cadastrar novos utilizadores que farão uso do ambiente computacional. Os novos elementos do sistema, foram divididos entre: **super usuário** e **usuário comum**. O primeiro tipo tem a possibilidade de executar todas as funções do sistema. Enquanto, o segundo tipo de usuário, executa somente a funcionalidade de geração de relatórios.
- **Geração de relatórios:** esta função busca apresentar para os gestores, dados gráficos a respeito da criação. Até o presente momento cinco diferentes gráficos foram elaborados, sendo eles: gráfico de evolução de alimentação, gráfico comparativo entre os frangos de corte, gráfico de busca de alimentação separado por períodos do dia, bem como, gráficos acerca de mudanças de temperatura e variações na umidade relativa do ar. Cada gráfico citado pode ser gerado escolhendo aviário, frango e datas de início e fim para obtenção dos dados.
- **Cadastro de frangos:** busca atender a etapa de entrada de novos frangos de corte a criação. O cadastro de novos elementos ao aviário acontece de forma automatizada, ou seja, cada ave recebe uma *tag* RFID que ao entrar em contato com o campo de leitura da antena, será lida e automaticamente alimentará o sistema.
- **Cadastro de aviários:** visa solucionar questões de cadastro de aviários junto ao sistema. O cadastramento de novos aviários acontece de forma manual, ou seja, sempre que uma nova produção for criada o **super usuário** do sistema, informará ao mesmo a sua criação, através do preenchimento de um formulário contido no próprio sistema.