



# Um estudo sobre o comportamento alimentar de frangos de corte utilizando a mineração de dados

A study on the feeding behavior of poultry using data mining

A. S. P. Rodrigues\*; E. N. Borges; R. Barwaldt

*Centro de Ciências Computacionais, Universidade Federal do Rio Grande, 96203-900, Rio Grande-RS, Brasil*

\* *alex\_sandrorodrigues@hotmail.com*

*(Recebido em 15 de maio de 2017; aceito em 22 de maio de 2017)*

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. Como forma de contribuir com este contexto, o presente trabalho buscou compreender a relação entre os períodos do dia e o tempo de vida dos frangos de corte no seu comportamento alimentar. Para que os resultados do presente trabalho fossem obtidos, foram adotados 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. Após o período de coleta de dados (20 dias) na Escola Técnica Estadual Santa Isabel, localizada na zona rural do município de São Lourenço do Sul, foi observado que os frangos que se encontram no estágio inicial de criação buscam se alimentar com maior frequência nos diversos períodos do dia, destacando-se, os períodos da noite e madrugada. Também se constatou que conforme os frangos crescem, a procura por alimentação diminuía de forma progressiva.

Palavras-chave: RFID, descoberta de conhecimento, frangos de corte

Poultry is an important sector of the Brazilian economy, accounting for 1.5% of the Brazilian GDP, as well as five million direct and indirect jobs. We notice the privileged Brazilian position in the world scenario, among the first three in the world, when it refers the production and consumption of the bird, and first when it refers to exports of the same. As a way of contributing to this context, the present work sought to understand the relationship between the daytime periods and the life time of broilers in their feeding behavior. In order to obtain the results of the present work, we adopted the automatic monitoring method, RFID, which can be seen as an identification technology that operates under radio waves, capable of sending and storing data remotely. Concomitant with the use of this technology, the discovery of knowledge in databases was used to discover implicit patterns in the datasets, transforming data without meaning into useful information. After the data collection period (20 days) at the Escola Técnica Estadual Santa Isabel, located in the rural area of the municipality of São Lourenço do Sul, it was observed that the chickens that are at the initial stage of breeding seek to feed more frequently in the different periods of the day, emphasizing some periods of the night and dawn. It was also found that as the chickens grew, the demand for food decreased progressively.

Keywords: RFID, knowledge discovery, poultry

## 1. INTRODUÇÃO

O setor de agronegócio é um dos mais competitivos do Brasil, pois representa 22% do produto interno bruto (PIB) nacional e 41% das exportações feitas pelo país. Destaca-se neste meio a avicultura que corresponde a 1,5% do PIB e 5 milhões de empregos diretos e indiretos, conforme dados de [1]. Estes dados também mostram que o país produziu mais de 12 milhões de toneladas de carne de frango, colocando o Brasil entre os três primeiros do mundo, quando se trata de produção e consumo da carne de frango, e em primeiro lugar quando se refere a exportações. A partir dos dados apresentados, fica evidente a importância do setor para a economia brasileira. No entanto, apesar de destacada posição neste âmbito, a busca por

melhorias deve ser constante, procurando saber como os fatores influenciam no desempenho e produtividade das cadeias produtivas avícolas.

Como forma de entender os fatores que influenciam o comportamento alimentar dos frangos de corte e, conseqüentemente, o desempenho da produção, o presente artigo, proveniente do projeto de extensão da Universidade Federal do Rio Grande (FURG) denominado "Tecnologia RFID como aliada a sustentabilidade agrícola" propõe a inserção de tecnologias de monitoramento e de informação no processo de acompanhamento de cadeia produtiva avícola, buscando: compreender a relação entre os períodos do dia, o tempo de vida dos frangos e o comportamento alimentar.

## 2. MATERIAL E MÉTODOS

Para realizar os objetivos propostos previamente, utilizou-se o método de identificação automática através de sinais de rádio, denominado *Radio Frequency Identification* (RFID), que pode ser visto como uma tecnologia de identificação que opera sob ondas de rádio, capaz de enviar e armazenar dados, além de, identificar, localizar e monitorar a posição de pessoas, animais ou outros objetos. Na figura 1 são mostrados os componentes do sistema RFID [2, 5].

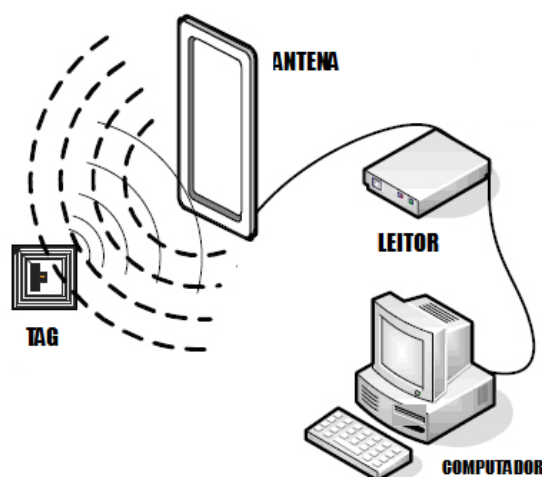


Figura 1: Componentes do RFID.

O sistema RFID é composto por três elementos, sendo eles: *tags*, antenas e leitores, cada qual, com uma determinada atribuição na estrutura. O primeiro componente citado, a *tag* tem como função identificar os diferentes elementos nos quais está acoplada. A antena é responsável por estabelecer a comunicação entre a *tag* e o leitor. Por fim, o leitor tem a tarefa de controlar acessos múltiplos, realizar a correção de erros, além de, enviar as informações coletadas pela antena para sistemas externos de processamento de dados.

Concomitante ao uso desta tecnologia, utilizou-se a descoberta de conhecimento em bases de dados, do inglês *Knowledge Discovery in Databases* (KDD). Conforme [7, 8, 9] este campo de estudo surgiu devido de ferramentas e técnicas tradicionais de análise de dados serem ineficazes com relação à avaliação de grandes volumes de dados, bem como, 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 [8], o KDD é a combinação de meios tradicionais de análise de dados junto com sofisticados algoritmos para avaliação de vastos conjuntos de dados. 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.

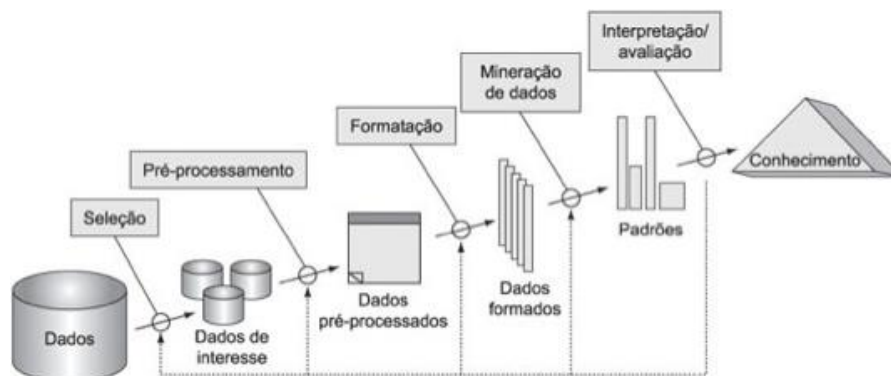


Figura 2: Processos da descoberta de conhecimento em base de dados.

Na figura 2 estão representados todos os processos envolvidos na descoberta de conhecimento em bases de dados.

As principais etapas do processo de KDD [4, 8, 9] serão resumidas a seguir. O 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.

A 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.

A interpretação 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.

Para o presente trabalho, a coleta de dados aconteceu na Escola Técnica Estadual Santa Isabel (ETESI), uma instituição de ensino agrotécnico localizada no interior do município de São Lourenço do Sul, durante 20 dias. Foi utilizada uma amostra de 15 frangos, cada qual, identificado por um registrador único, armazenado num banco de dados. Na figura 3, pode-se observar o local disponibilizado para a coleta de dados.



Figura 3: Local disponibilizado para coleta de dados.

Como dito anteriormente, o principal objetivo do estudo foi analisar o comportamento alimentar relacionando o tempo de vida e os períodos do dia na alimentação dos frangos. Desse modo, optou-se por instalar o sistema RFID diretamente na área de alimentação dos animais. Assim, 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 um sistema de informação responsável por armazenar os dados coletados pelo RFID. Deve-se ressaltar o tipo de material da estação de alimentação, inicialmente o metal, no entanto, testes preliminares revelaram uma alta taxa de interferência do material nas ondas de radiofrequência emitidas pelo RFID, por esse motivo, adotaram-se comedouros de madeira, visto que, o material quando combinado com o material orgânico, apresentou taxas de interferência menores.

Depois de concluída a etapa de coleta dos dados, iniciou-se a fase de análise dos dados, que procurou responder, por exemplo, um frango que está no 20º dia de vida, alimenta-se com frequência? Se sim, em qual período do dia? Para responder estas questões, utilizou-se a descoberta de conhecimento em base de dados.

Na tabela 1, pode-se observar um fragmento da tabela onde estão mantidos os dados, antes da etapa de pré-processamento.

*Tabela 1: Organização dos dados antes do pré-processamento.*

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

A coluna *id\_frango* representa o número de identificação de cada frango no sistema. A coluna *zona\_leitura*, indica se o frango encontra-se na zona de alimentação para a esta coluna, tem-se dois possíveis valores, *true* ou *false*, onde o primeiro representa que o frango localiza-se na zona de alimentação, o segundo valor indica o oposto. As demais colunas mostram o dia e a hora que o frango situava-se na área de alimentação.

No pré-processamento, buscou-se transformar os dados contínuos em dados categóricos, dessa maneira, as colunas que se encontravam dados referentes aos horários e datas de entrada e saída dos frangos na área de alimentação, foram separadas em períodos do dia e fases de crescimento. Também foi considerado um tempo mínimo de 15 segundos de permanência na zona de leitura, desse modo, múltiplas leituras que estavam no intervalo de tempo citado foram desconsideradas. Na Tabela 2 e Tabela 3 são apresentados como dados foram organizados para posterior análise.

*Tabela 2: Períodos de análise.*

<b>Período</b>	<b>Início</b>	<b>Fim</b>
<b>Madrugada</b>	00:00:00	05:59:59
<b>Manhã</b>	06:00:00	11:59:59
<b>Tarde</b>	12:00:00	17:59:59
<b>Noite</b>	18:00:00	23:59:59

Tabela 3: Fases de análise.

<b>Fase</b>	<b>Início</b>	<b>Fim</b>
<b>1</b>	19/11/2015	26/11/2015
<b>2</b>	27/11/2015	04/12/2015
<b>3</b>	05/12/2015	11/12/2015

Na tabela 4 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 se refere a presença dos frangos na zona de alimentação foi deslocada para a última posição transformando-se no atributo alvo.

Tabela 4: Organização dos dados após o pré-processamento.

<b>Fase</b>	<b>Período</b>	<b>Zona de Alimentação</b>
<b>1</b>	Madrugada	T
<b>1</b>	Madrugada	F
<b>1</b>	Madrugada	T
<b>2</b>	Manhã	F
<b>2</b>	Manhã	T
<b>2</b>	Tarde	F
<b>3</b>	Tarde	T
<b>3</b>	Noite	F
<b>3</b>	Noite	T

Após a etapa de pré-processamento, os dados (46.506 instâncias) transformados foram usados como entrada para o software de mineração de dados Weka, onde a tarefa de mineração escolhida foi a classificação de dados que consiste em 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 [3].

Para elaborar o modelo pretendido pela classificação escolheu-se o algoritmo J48, que surgiu da necessidade de recodificação do algoritmo C4.5, originalmente escrito na linguagem C para Java [9].

O algoritmo J48 é baseado em árvores de decisão, que conforme [3], “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”. A configuração adotada para a execução do algoritmo seguiu os padrões pré-definidos pelo Weka, tais como: fator de confiança de 0.25, validação cruzada (*cross-validation*) de 10 partições.

### 3. RESULTADOS E DISCUSSÃO

Na figura 4 é possível observar o modelo gerado pelo algoritmo J48 com taxa de acurácia de 72,436%.

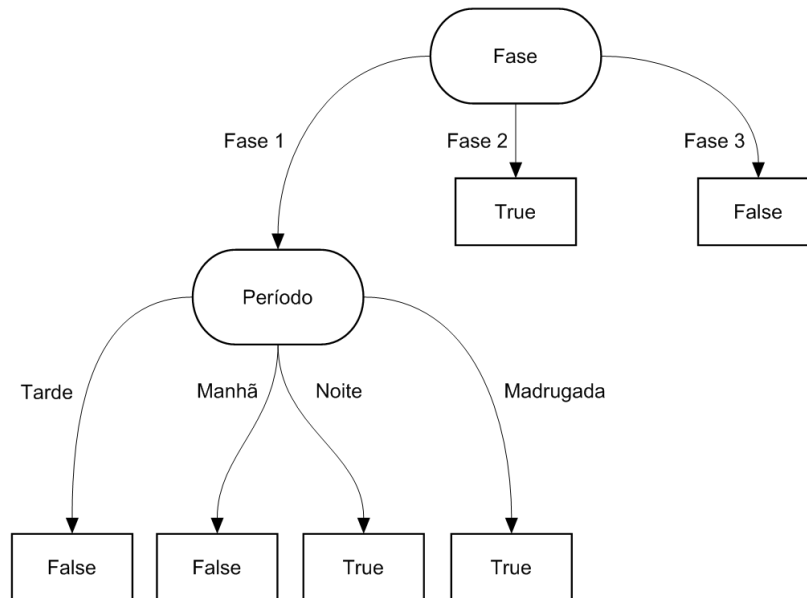


Figura 4: Modelo induzido pelo algoritmo C4.5.

Outras medidas de avaliação do modelo, também devem ser levadas em consideração, visto que, contribuem para a validade do modelo gerado. A primeira medida de avaliação a ser considerada é a precisão (*Precision*), na qual, mede a porcentagem de acerto entre as instâncias preditas com um determinado rótulo. Para a classe *true* que significa a busca dos frangos por alimentação, a taxa de precisão foi de 69,6%, enquanto para classe *false* que representa a não procura de alimentação por parte das aves, a medida de avaliação apresentou 75,8%.

A segunda métrica trata-se da revocação (*Recall*), onde o objetivo é medir a porcentagem de acerto entre as instâncias de uma determinada classe. Para o rótulo, *true* a taxa de revocação foi de 65,9%, enquanto para o rótulo *false* foi de 78,8%. A última medida ponderada no presente trabalho trata-se da Medida F (*F-Measure*) que refere-se a média harmônica entre a precisão e a revocação. A classe *true* apontou valor de 70,5%, já a classe *false* indicou o 73,9%.

Por fim, é observado que os frangos que estão na fase 2 buscam alimentar-se com maior regularidade, isto é, percebido pelo fato de a árvore de decisão indicar que os mesmos, passam a maior parte do tempo dentro da zona de alimentação.

O contrário acontece com os frangos que estão na fase 3, onde detectou-se que conforme os frangos cresciam, a procura por alimentação diminuía de forma progressiva. Também se pode perceber que os frangos que estão na fase 1, alimentam-se em diversos períodos, sendo, os períodos da noite e madrugada, os que apresentam maior procura.

Na figura 5, onde o eixo y contém dados referentes a quantidade de vezes que o frango adentrou a zona de leitura do RFID, enquanto o eixo x armazena dados relativos aos dias de supervisionamento, é possível confirmar as informações descobertas pelo processo de DCBD, onde o frango que esta na fase 1 de monitoramento apresenta uma procura por alimentação crescente. Enquanto, os frangos que se encontram nas fases 2 e 3 de acompanhamento mostram uma crescente queda na busca por alimentação.

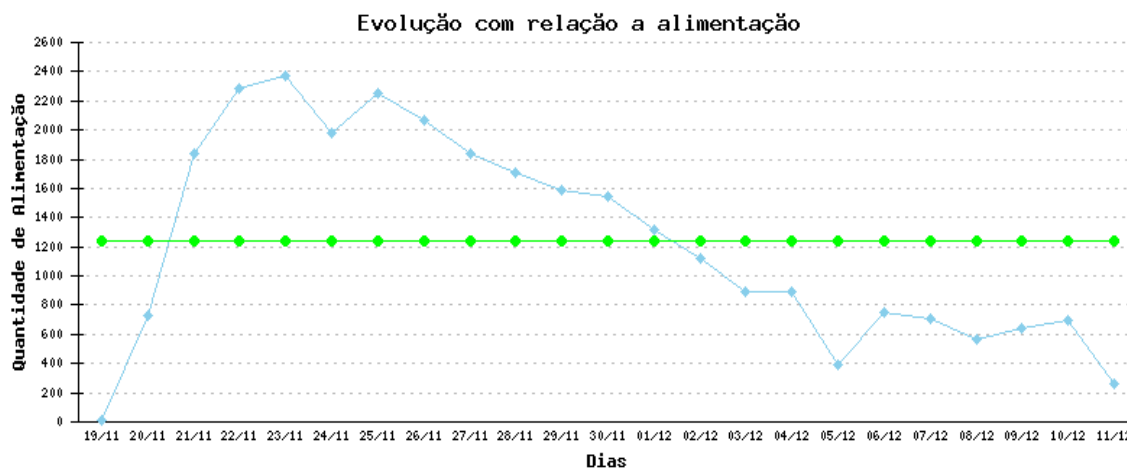


Figura 5: Evolução da alimentação no tempo.

#### 4. CONCLUSÃO

No presente artigo, foi possível observar que o RFID não gerou interferência no desenvolvimento dos frangos, mostrando que a tecnologia é factível de ser utilizada no processo de monitoramento de cadeias produtivas avícolas. Também foi possível verificar que, a partir dos dados coletados pelo RFID, a análise por descoberta de conhecimento gerou conclusões interessantes sobre o comportamento da cadeia produtiva, que posteriormente podem ser usadas na criação e desenvolvimento dos frangos de corte. No entanto, é importante destacar que apesar dos resultados satisfatórios e reconhecidos por um especialista em avicultura da instituição ETESI, é requerida uma quantidade maior de validações, para provar que as informações obtidas neste trabalho evidenciam uma realidade em termos de comportamento dos frangos e não uma anormalidade.

Salienta-se que o trabalho teve importante relevância no que tange a necessidade de aperfeiçoamento desta área de pesquisa na região e arredores, onde há concentração de produtores rurais, engajados através das cooperativas. Por fim, espera-se aprimorar demais técnicas, utilizando estas ferramentas em cadeias produtivas de aves, definir novas metodologias de análise de dados, bem como, inserir o monitoramento de variáveis de temperatura, umidade relativa do ar e umidade da cama do aviário, além da possibilidade de monitorar o peso dos frangos de corte.

#### 5. REFERÊNCIAS BIBLIOGRÁFICAS

1. ABPA. Relatório anual [Internet]; 2016 [cited 2017 May 10]. Available from: <http://goo.gl/4Mg09a>.
2. Domdouzis K, Kumar B, Anumba C. Radio-frequency identification (rfid) applications: a brief introduction. *Advanced Engineering Informatics* 2007; 21(4): 350-355, doi: 10.1016/j.aei.2006.09.001.
3. Faceli K. Inteligência artificial: uma abordagem de aprendizado de máquina. Rio de Janeiro: LTC; 2011.
4. Fayyad UM, Piatetsky-Shapiro G., Smyth P, Uthurusamy R. *Advances in knowledge discovery and data mining*. Menlo Parck: American association for artificial intelligence; 1996.
5. Gomes HMC. Construção de um sistema de rfid com fins de localização especiais [Dissertation]. Universidade do Porto; 2007.
6. Mendes AA. Panorama da avicultura nacional e perspectivas do setor [Internet]; 2014 [cited 2017 May 10]. Available from: <http://goo.gl/isEYYb>.
7. Silva MPS. Mineração de dados: conceitos, aplicações e experimentos com weka. Porto Alegre: Sociedade Brasileira de Computação; 2004.
8. Tan PN, Steinbach M, Kumar V. *Introdução ao data mining: mineração de dados*. Rio de Janeiro: Ciência Moderna; 2009.
9. Witten IH, Frank E. *Data mining: practical machine learning tools and techniques*. San Francisco: Morgan Kaufmann Publishers; 2011.