

UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ZOOTECNIA

**HADJA LORENA RANGEL UCHÔA CAVALCANTI DE MENEZES
COSTA**

**ESTUDO PRELIMINAR PARA PREDIÇÃO DE CONSUMO DE NUTRIENTES E
DESEMPENHO DE OVINOS DESLANADOS EM CONFINAMENTO BASEADO EM
APRENDIZAGEM DE MÁQUINA**

RECIFE

2023

**HADJA LORENA RANGEL UCHÔA CAVALCANTI DE MENEZES
COSTA**

**ESTUDO PRELIMINAR PARA PREDIÇÃO DE CONSUMO DE
NUTRIENTES E DESEMPENHO DE OVINOS DESLANADOS EM
CONFINAMENTO BASEADO EM APRENDIZAGEM DE MÁQUINA**

Dissertação apresentada ao Programa de
Pós-Graduação em Zootecnia da
Universidade Federal Rural de
Pernambuco para obtenção do título de
Mestre em Zootecnia

Área de concentração: Zootecnia

Orientador: Prof.^a Dra. Antonia
Sherlânea Chaves Vêras

Coorientador(es):
Prof. Dr. Maxwell Guimarães de
Oliveira
Dr. Jasiel Santos de Morais

**RECIFE
2023**

Dados Internacionais de Catalogação na Publicação
Universidade Federal Rural de Pernambuco
Sistema Integrado de Bibliotecas
Gerada automaticamente, mediante os dados fornecidos pelo(a) autor(a)

- C837e Costa, Hadja Lorena Rangel Uchôa Cavalcanti de Menezes
Estudo preliminar para predição de consumo de nutrientes e desempenho de ovinos deslanados em confinamento baseado em aprendizagem de máquina / Hadja Lorena Rangel Uchôa Cavalcanti de Menezes Costa. - 2023.
35 f. : il.
- Orientadora: Antonia Sherlanea Chaves Veras.
Coorientador: Maxwell Guimaraes de Oliveira.
Inclui referências.
- Dissertação (Mestrado) - Universidade Federal Rural de Pernambuco, Programa de Pós-Graduação em Zootecnia, Recife, 2023.
1. Aprendizado de máquina. 2. Inteligência artificial. 3. RNA. 4. Santa Inês. I. Veras, Antonia Sherlanea Chaves, orient. II. Oliveira, Maxwell Guimaraes de, coorient. III. Título



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ZOOTECNIA

**ESTUDO PRELIMINAR PARA PREDIÇÃO DE CONSUMO DE
NUTRIENTES E DESEMPENHO DE OVINOS DESLANADOS EM
CONFINAMENTO BASEADO EM APRENDIZAGEM DE MÁQUINA**

Dissertação elaborada por
HADJA LORENA RANGEL UCHÔA CAVALCANTI DE MENEZES COSTA

Aprovado em/...../.....

BANCA EXAMINADORA

Prof^ª. Dra. Antonia Sherlânea Chaves Vêras
Universidade Federal Rural de Pernambuco

Prof. Dr. Cláudio Elisio Calazans Campelo
Universidade Federal de Campina Grande

Prof^ª. Dra. Renata Valéria Regis de Sousa Gomes
Universidade Federal Rural de Pernambuco

RESUMO

Como o principal motivo do confinamento de animais de corte é produzir carne com qualidade para abastecer o mercado interno com oferta regular e ao menor custo de produção, utiliza-se assim, um sistema baseado em Redes Neurais Artificiais (RNA) que permita a previsão do desempenho animal, com base em entradas iniciais e séries históricas. Parece promissora, permitindo com que animais de baixo desempenho sejam identificados previamente e descartados de maneira a maximizar a produção. Dada a importância dessa informação, bem como do consumo de nutrientes pelos animais, o presente estudo tem por objetivo investigar o uso de RNA como preditores do consumo de nutrientes e, a partir da composição química da dieta, prever o desempenho de ovinos deslançados em confinamento. O banco de dados foi obtido a partir das informações do acervo de experimentos científicos realizados no setor de ovinos do Departamento de Zootecnia da Universidade Federal Rural de Pernambuco. Para testar a eficácia da RNA em relação à previsão das características desejadas foram exploradas outras nove técnicas de previsão com características distintas, com o intuito de avaliar diferentes cenários e observar qual a melhor técnica para cada situação. Os resultados mostraram que a RNA apresentou um desempenho abaixo do esperado, quando comparada a outras técnicas. Isso pode ser uma consequência do baixo volume de dados disponíveis para treinamento e validação dos modelos de previsão. Novos estudos são necessários para explorar outras ferramentas de previsão que se apresentam promissoras para a ciência animal como um todo. Em resumo, embora as RNAs tenham capacidade técnica promissora na área da ciência animal, é importante avaliar cuidadosamente a seleção de variáveis e técnicas adequadas para cada caso específico, visando obter resultados mais precisos e confiáveis.

Palavras-chave: Aprendizado de Máquina, Inteligência Artificial, RNA, Santa Inês.

ABSTRACT

Once the main reason for confinement of livestock animals is to produce quality meat to supply the domestic market with regular supply and at the lowest production cost, the use of an Artificial Neural Network (ANN) system that allows the prediction of animal performance based on initial inputs and historical data seems promising, enabling low-performing animals to be identified in advance and discarded to maximize production. Given the importance of this information, as well as nutrient consumption by animals, the present study aims to investigate the use of ANN as predictors of nutrient intake and, from the chemical composition of the diet, predict the performance of shorn sheep in confinement. The database was obtained from information from scientific experiments conducted in the sheep sector of the Department of Animal Science at the Federal Rural University of Pernambuco. To test the efficacy of ANN in predicting desired characteristics, nine other prediction techniques with distinct features were explored to evaluate different scenarios and observe the best technique for each situation. Dealing with all predicted points, such as average daily intake, average daily gain, and final body weight, the ANN demonstrated a performance well below expectations when compared to other techniques included in the analysis. This may be a consequence of the low volume of data available for training, verification, and checking of the prediction models. Further studies are needed to explore other promising prediction tools for animal science as a whole. In summary, although ANNs may be a promising technique in the field of animal science, it is important to carefully evaluate the selection of variables and appropriate techniques for each specific case to obtain more accurate and reliable results.

Keywords: Machine Learning, Artificial Intelligence, ANN, Santa Inês.

LISTA DE ILUSTRAÇÕES

	Página
Figura 1. Exemplo de validação cruzada estratificada 5-fold.....	23
Figura 2. Gráfico de barras referentes à acurácia, precisão, recall e ao F1-score das técnicas de predição do GMD.....	24
Figura 3. Gráfico de barras referentes à acurácia, precisão, recall e ao F1 Score das técnicas de predição do CMS.....	25
Figura 4. Gráfico de barras referentes à acurácia, precisão, recall e ao F1 Score das técnicas de predição do PCf.....	26

LISTA DE QUADROS

	Página
Quadro 1. Equações utilizadas para estimativa de variáveis não informadas pelos autores...15	
Quadro 2. Tabela de variáveis utilizadas nos modelos de predição utilizando AM.....19	

SUMÁRIO

	Página
1. INTRODUÇÃO	9
2. REVISÃO DE LITERATURA	11
3. MATERIAL E MÉTODOS	13
3.1 Coleta de dados e critério de seleção.....	13
3.2 Processamento dos dados.....	15
3.3 Seleção das técnicas e do conjunto de características a serem testadas.....	17
3.4 Ferramentas utilizadas e desenvolvimento da RNA.....	20
4. RESULTADOS E DISCUSSÃO	23
5. CONCLUSÃO	28
6. REFERÊNCIAS	30

1. INTRODUÇÃO

O desempenho animal é reflexo da ingestão de alimentos e, portanto, a partir do conhecimento do consumo de matéria seca (MS) e de seus constituintes, é possível inferir sobre a resposta animal.

Uma forma de otimizar e melhorar a acurácia desse tipo de avaliação seria fazer uso de Inteligência Artificial (IA). A IA pode ser explicada, de maneira mais simples, como um algoritmo capaz de sistematizar a resolução de problemas, tendo como ponto de partida a observação de dados oriundos dos complexos sistemas existentes na natureza, permitindo a execução de tarefas complexas como o reconhecimento de padrões em imagens e sons, processamento de linguagem e planejamento e predição de ações. Sendo assim, uma IA tem como principal objetivo desenvolver modelos computacionais de como os organismos vivos se comportam diante de problemas e se adaptam às diferentes situações (COSTA, 2009).

O uso de IA pode auxiliar nas atividades pecuárias, como monitoramento e controle dos rebanhos, principalmente no âmbito da análise de dados e tomada de decisões dentro das propriedades rurais, possibilitando uma análise mais acurada das mensurações de peso obtidos dos animais (SANT'ANA et al., 2021), auxiliando o responsável na tomada decisões de maneira mais rápida e eficiente, com base na realidade da produção animal.

Diante deste contexto, os métodos baseados em Redes Neurais Artificiais (RNA) parecem ser, particularmente, apropriados para predição das exigências nutricionais dos animais, em que o conhecimento da ingestão de alimentos e do desempenho animal se destacam; por sua capacidade de prever resultados de problemas, por meio de levantamento de conjuntos de dados históricos dos problemas sem conhecer as interações entre os parâmetros, ainda que as interações sejam altamente não lineares (ZURADA, 1992).

As RNA são um tipo de Aprendizado de Máquina (AM), um ramo da área de IA, que vem atraindo muita atenção devido à sua grande capacidade de autoaprendizagem e boa precisão no mapeamento de relações não lineares complexas; além de apresentarem excelente desempenho com grandes volumes de dados e possuírem bibliotecas avançadas de código aberto (LIU et al., 2021; XU et al., 2021).

A utilização dessa técnica na prática permite a rápida resolução de problemas, sem a necessidade de definição de lista de regras ou de modelos matemáticos precisos, como ocorre

muitas vezes com modelos de regressões lineares múltiplas e outros comumente usados. As RNA utilizam programas de instruções sequenciais, utilizando regras de aprendizagem que são obtidas através de exemplos, ou seja, dos dados disponíveis para essa aprendizagem, permitindo adquirir poder de generalização suficiente para reconhecer padrões e prever cenários (BEZERRA et al., 2020).

São diversos os potenciais de aplicação das RNA na produção animal, atuando como controle dos elementos meteorológicos em instalações zootécnicas como resposta ao comportamento postural, à vocalização e o consumo de energia metabolizável; a adequação do manejo com base nas variáveis fisiológicas, nos aspectos nutricionais e na resposta produtiva; a caracterização da sanidade do rebanho sem a utilização de práticas invasivas; e a seleção e melhoramento genético e o monitoramento de operações de pré-abate na determinação da qualidade do produto final, apenas algumas das aplicações dessa ferramenta (PANDORFI et al., 2011).

Como o principal motivo do confinamento de animais de corte é produzir carne com qualidade e com oferta regular para abastecer ao mercado interno e ao menor custo de produção, a utilização de um sistema baseado em RNA que permita a previsão do desempenho dos animais com base em entradas iniciais e séries históricas parece promissora, permitindo com que animais de baixo desempenho sejam identificados previamente e descartados de maneira a maximizar a produção. Dada a importância dessa informação, bem como da quantidade e qualidade de consumo de nutrientes pelos animais, o presente estudo tem por objetivo investigar o uso de RNA como preditores do consumo de matéria seca e seus constituintes e, a partir da composição química da dieta, prever o desempenho de ovinos deslançados em confinamento de forma automatizada.

Além disso, as RNA podem ser usadas para prever o consumo de matéria seca e seus constituintes em ovinos e, a partir da composição química da dieta, prever seu desempenho em confinamento. Isso pode permitir que o produtor ajuste a dieta dos animais para maximizar o desempenho e reduzir custos, tornando a produção mais eficiente e sustentável. A utilização de RNA nesse contexto também pode auxiliar na identificação de possíveis problemas de saúde ou nutrição nos animais, permitindo uma intervenção mais precoce e eficiente.

Com base no conhecimento de que o desempenho animal pode ser inferido a partir do

conhecimento do consumo de matéria seca e de seus constituintes, uma hipótese a ser considerada é que as RNA são uma boa alternativa para prever o desempenho de ovinos em confinamento, a partir da composição química da dieta. Assim, a hipótese nula (H_0) poderia ser formulada como "as RNA são uma boa preditora para o desempenho de ovinos em confinamento, a partir da composição química da dieta" e a hipótese alternativa (H_a) como "as RNA não são uma boa alternativa para prever o desempenho de ovinos em confinamento, a partir da composição química da dieta". Essas hipóteses poderiam ser testadas por meio de análises estatísticas apropriadas, utilizando-se conjuntos de dados históricos e amostras de ovinos em confinamento, com variações na composição química da dieta.

2. REVISÃO DE LITERATURA

As recomendações nutricionais para ovinos mais usadas no Brasil são oriundas dos comitês americano, National Research Council (NRC), e o Agricultural and Food Research Council (AFRC) do Reino Unido. Contudo, há necessidade de pesquisas para ajuste das exigências nutricionais de ovinos de genótipos diferentes, nas diversas condições climáticas de criações; sobretudo, em relação à necessidade energética e, preferencialmente por meio de métodos que não sejam laboriosos e que não comprometam as carcaças dos animais. Adicionalmente, a detecção precoce de animais de baixo desempenho também é importante para decisões oportunas de abate com seus benefícios econômicos associados, uma vez que diminui os custos com alimentação. Além do mais, a previsão do desempenho é um fator importante para tomada de decisões com vistas à oferta regular de produtos no mercado.

Oliveira et al. (2017), em uma meta-análise para prever as exigências nutricionais de ovinos criados na região Nordeste do Brasil, observaram que os sistemas internacionais, como o NRC e o AFRC, superestimam as exigências nutricionais para as raças e genótipos de ovinos brasileiros, especialmente as energéticas e proteicas. Quando utilizaram dados de animais com peso corporal (PC) de 20 kg e ganhos médios de 100 g/dia, verificaram que os animais necessitam de 1.337 Mcal/dia de energia metabolizável (EM); Enquanto que, com base nos sistemas americano e o britânico, esses requerimentos seriam de 1.505 e 1.600 Mcal/dia, respectivamente.

Assim, para um confinamento com duração de 60 dias, usando as equações sugeridas pelos referidos autores, o produtor teria uma economia de 3,21 e 5,06 kg de milho por animal. Vale ressaltar que, para compor o banco de dados desta meta-análise, os pesquisadores usaram

os dados brutos de experimentos, com o objetivo de melhorar a acurácia dos resultados, devido ao maior número de repetições; sendo que 61,32% das informações utilizadas originaram-se de experimentos com ovinos da raça Santa Inês, devido à maior expressão dessa raça no rebanho da região Nordeste do Brasil, devido às suas características adaptativas às condições climáticas locais, além de apresentar bom desempenho produtivo e reprodutivo. Alguns autores que mencionam a importância da raça Santa Inês na região Nordeste são Lima et al. (2014) e Silva et al. (2018).

Uma forma de otimizar e melhorar a acurácia desse tipo de avaliação seria fazer uso de Inteligência Artificial (IA). O uso de IA pode auxiliar nas atividades pecuárias, como monitoramento e controle dos rebanhos, principalmente no âmbito da análise de dados e tomada de decisões dentro das propriedades rurais, possibilitando uma análise mais acurada das mensurações de peso obtidos dos animais (SANT'ANA et al., 2021) e, assim, viabilizando a tomada de decisões de maneira mais rápida e eficiente pelos produtores, com base na realidade da produção animal.

Embora pareça uma tecnologia nova no ramo da pecuária, muitos estudos já foram realizados e têm demonstrado a eficácia dessa tecnologia nos sistemas de produção animal, suas aplicações e as possibilidades a serem exploradas com o auxílio desse tipo de monitoramento; como é o caso do trabalho de Sant'Ana et al. (2021), que investigaram a utilização de técnicas de Visão Computacional na avaliação do peso vivo de ovinos; Bezen, Edan e Halachmi (2020) abordaram o mesmo tema, dessa vez voltados para ingestão de alimentos em vacas; Samperio et al. (2021) avaliaram a estimativa de peso vivo de cordeiro utilizando a análise computacional de imagens 3D, entre tantos outros trabalhos abordando a introdução dessa nova ferramenta no auxílio do gerenciamento de rebanhos.

Nesse contexto, os métodos baseados em Redes Neurais Artificiais (RNA) parecem ser, particularmente, apropriados para predição das exigências nutricionais dos animais, em que o conhecimento da ingestão de alimentos e do desempenho animal se destacam, por sua capacidade de prever resultados de problemas, por meio de levantamento de conjuntos de dados históricos dos problemas sem conhecer as interações entre os parâmetros, ainda que as interações sejam altamente não lineares (ZURADA, 1992).

As RNA são um tipo de aprendizado de máquina, uma vertente da área da Inteligência Artificial, que vem atraindo muita atenção devido a sua grande capacidade de autoaprendizagem e boa precisão no mapeamento de relações não lineares complexas; além

disso, também de apresentarem excelente desempenho com dados crescentes e permitirem bibliotecas avançadas de código aberto (LIU et al., 2021; XU et al., 2021).

No uso de RNA, a priori, não há necessidade de começar com um modelo, nem de identificar previamente todas as variáveis necessárias. As Redes Neurais têm a facilidade de combinação de diferentes tipos de entrada (binários e contínuos) e são potencialmente vantajosas na modelagem de processos biológicos frequentemente caracterizados como altamente não lineares (LACROIX et al., 1995).

Tal técnica tem sido aplicada em muitos casos com sucesso, mas relativamente com pouca pesquisa no âmbito da produção animal, quando comparado com outras áreas que já empregam essa tecnologia há mais tempo. Antes as pesquisas realizadas se concentraram principalmente na pecuária leiteira, no que diz respeito à previsão da produção individual de leite, gordura e proteína e detecção de doenças, bem como no comportamento animal e reconhecimento facial (KOMINAKIS et al., 2002), ou sobre estresse térmico em bovinos de corte (SOUSA et al., 2018).

No entanto, atualmente é perceptível uma crescente expansão para outras criações animais como também para outros tipos de análises, como é o caso dos trabalhos de Sant'Ana et al. de 2021, Bezen, Edan e Halachmi de 2020, Samperio et al. em 2021 já mencionados anteriormente, além de outros, como os trabalhos de Shadpour et al. (2022) e Alcantara et al. (2022) que mesmo trabalhando com gado leiteiro, abordam novas área de utilização, sendo o primeiro na predição de ingestão de matéria seca e o segundo mais voltado para a parte de protocolos de reprodução.

Os trabalhos envolvendo ovinos ainda são relativamente recentes e escassos; no entanto, exemplos como o de Bezerra et al. (2020), que trabalham com RNA na predição de estresse em ovelhas prenhas, Sant'Ana et al. de 2021 e Samperio et al. em 2021 já mencionados anteriormente, assim como o trabalho de Silva (2021) utilizando RNA para predição de estresse em ovelhas prenhas e lactantes. Todavia, trabalhos relevantes envolvendo redes neurais artificiais para o desempenho de ovinos não foram encontrados, de maneira que o presente estudo preliminar poderá auxiliar em pesquisas futuras, além de contribuir de maneira significativa com o banco de dados criado.

3. MATERIAL E MÉTODOS

3.1 Coleta de dados e critério de seleção

O banco de dados utilizado foi obtido a partir dos dados históricos do acervo de

experimentos científicos (teses e dissertações defendidas junto ao Programa de Pós-Graduação em Zootecnia) realizados no setor de ovinos do Departamento de Zootecnia da Universidade Federal Rural de Pernambuco. Para o estudo preliminar, foram utilizadas médias de tratamentos de 1.675 animais, obtidos de 49 trabalhos coletados no acervo digital do Programa e 26 do acervo físico, somando um total de 75 trabalhos coletados.

Como critério de seleção dos trabalhos, foram consideradas as informações provenientes de dissertações e teses que utilizaram ovinos Santa Inês e seus mestiços, machos não castrados. As seguintes variáveis foram consideradas na coleta: composição química das dietas e consumos de matéria seca (MS), matéria orgânica (MO), proteína bruta (PB), extrato etéreo (EE), fibra em detergente neutro (FDN), fibra em detergente ácido (FDA), carboidratos não fibrosos (CNF), nutrientes digestíveis totais (NDT); além de dados de desempenho: peso corporal inicial (PCi), peso corporal final (PCf), ganho em peso médio diário (GMD), pesos e rendimentos de carcaças (quente e/ou fria) (RCQ/RCF), conversão alimentar (CA), eficiência alimentar (EA), rendimento biológico (RB) e perdas por resfriamento (PR) da carcaça.

Os dados coletados foram tabulados em planilhas no Microsoft Excel, totalizando 197 registros (ou linhas). Após a etapa de tabulação, foi efetuado o pré-processamento e a organização dos dados, onde trabalhos que apresentavam informações insuficientes a respeito da dieta ou do animal, e cujas variáveis não foi possível estimar, foram retiradas da base de dados, causando redução no número de linhas disponíveis para treinamento dos modelos. Também foram removidos trabalhos que utilizaram fêmeas e animais que não fossem da raça Santa Inês, ou, mestiços. Trabalhos que eram complementares, ou seja, trabalhos de diferentes autores que dividiram uma pesquisa em duas, foram associados. Ao final dessa etapa, a planilha contava com 111 médias de tratamentos de 940 animais obtidos a partir de 32 trabalhos.

É importante frisar que o banco de dados criado apresenta grande importância, pois, permite o acesso a informações valiosas sobre o desempenho de ovinos Santa Inês e seus mestiços em diferentes experimentos científicos, especialmente em relação à nutrição e ao desempenho produtivo. A partir dessas informações, é possível obter insights sobre a melhor forma de alimentação e manejo desses animais, além de possibilitar a realização de novos estudos e a comparação com resultados de outras pesquisas. Essa base de dados pode ser útil tanto para pesquisadores quanto para profissionais que trabalham com ovinocultura, permitindo a tomada de decisões embasadas em dados científicos confiáveis.

3.2 Pré-processamento dos dados

Em seguida, foi realizada a etapa de pré-processamento dos dados. Para os trabalhos que não possuíam informações como concentrações e consumos de NDT e PB, tais variáveis foram calculadas ou estimadas. Além disso, os PCi, PCf, CA, EA, PC ao abate/jejum, PCV, RCQ, RCF, RB e PR foram estimados utilizando equações de predição específicas para cada uma dessas variáveis avaliadas, como mostra o Quadro 1.

Quadro 1. Equações utilizadas para estimativa de variáveis não informadas pelos autores.

Variável predita	Equação utilizada	Fontes das informações utilizadas para estimativas
CNDT	$CMS \times (NDT \text{ dieta}/1000)$	Santos Filho (2012) Ferraz (2013) Alves (2012)
CNDTe	$(- 3,84 + 1,064 \times (DMS/10)) \% \times CMS$	Andrade (2010) Mattos (2009)
PCi	$PCf - ((GMD/1000) \times \text{Dias})$	Xenofonte (2006)
PCfe	$(PC \text{ abate/jejum} + 0,547)/0,9313$	Freitas (2013) Oliveira (2018) Oliveira (2013) Alves (2002) Vasconcelos (2008)/Souto Maior (2008)
PCf	$((GMD \times \text{Dias})/1000) + PCi$	Andrade (2011)/Urbano (2011) Alves (2012)
CA	CMS/GMD^a	Andrade (2010) Oliveira (2013) Alves (2012)/Xavier (2012)
EA	GMD/CMS^a	Silva (2017)/Santos (2017) Cunha (2006) Silva (2012)/Urbano (2015) Aguiar (2007) Andrade (2010) Vasconcelos (2013) Santos Filho (2012) Oliveira (2013) Ferraz (2013) Moura (2013) Maciel (2012) Andrade (2011) Alves (2012)

		Alves (2002) Xenofonte (2006) Silva (2012) Bezerra (2015) Alves (2012) Andrade (2010) Mattos (2009) Coelho (2014) Silva (2015) Souza Júnior (2007) Lopes (2019) Vasconcelos (2008)/Souto Maior (2008)
PC abate/jejum	$- 0,547 + (0,9313 \times PCf)^b$	Andrade (2010) Santos Filho (2012) Ferraz (2013) Souza Júnior (2007)
PCV	$- 1,4944 + (0,8816 \times PC \text{ abate/jejum})^b$	Aguiar (2007) Andrade (2010) Andrade (2010) Mattos (2009) Coelho (2014) Souza Júnior (2007)
RCQ	$(PCQ/PC \text{ abate/jejum}) \times 100^c$	Andrade (2010) Andrade (2011)/Urbano (2011)
RCF	$(PCF/PC \text{ abate/jejum}) \times 100^c$	Andrade (2010) Oliveira (2018) Andrade (2011)/Urbano (2011)
RB	$(PCQ/PCV) \times 100^d$	Freitas (2013) Aguiar (2007) Andrade (2010) Andrade (2011)/Urbano (2011) Mattos (2009) Coelho (2014)
PR	$((PCQ - PCF)/PCQ) \times 100^c$	Freitas (2013) Andrade (2011)/Urbano (2011) Alves (2012) Silva (2015)

CNDT: Consumo de Nutrientes Digestíveis Totais (g/dia); **CNDTe**: Consumo de Nutrientes Digestíveis Totais Estimado (g/dia); **CMS**: Consumo de Matéria Seca (g/dia); **NDT dieta**: Nutrientes Digestíveis Totais da Dieta (g/kg); **DMS**: Digestibilidade da Matéria Seca (g/kg); **PCi**: Peso Corporal Inicial (kg); **PCfe**: Peso Corporal Final Estimado (kg); **PCf**: Peso Corporal Final (kg); **GMD**: Ganho Médio Diário (g/dia); **Dias**: Dias de Experimento (dias); **CA**: Conversão Alimentar; **EA**: Eficiência Alimentar; **PC abate/jejum**: Peso Corporal ao Abate/Jejum (kg); **PCV**: Peso de Corpo Vazio (kg);

RCQ: Rendimento de Carcaça Quente (%); **PCQ:** Peso de Carcaça Quente (kg); **RCF:** Rendimento de Carcaça Fria (%); **PCF:** Peso de Carcaça Fria (kg); **RB:** Rendimento Biológico da Carcaça (%); **PR:** Perdas por Resfriamento da Carcaça (%). ^aFernandes et al. (2014), ^bHerbster et al. (2020), ^cOliveira et al. (2002), ^dPereira (2016).

3.3 Seleção das técnicas e do conjunto de características a serem testadas

A presente pesquisa foi realizada em parceria com o Centro de Ciências da Computação da Universidade Federal de Campina Grande, e por meio de um aluno de Programa Institucional de Bolsas de Iniciação Científica (PIBIC) que desenvolveu a programação e auxiliou durante todo o processo das análises dos dados e construção dos preditores.

Para testar a eficácia da RNA em relação à predição das características desejadas, foram utilizadas outras nove técnicas de predição em AM, com características distintas entre si, com o intuito de avaliar diferentes cenários e observar qual a melhor técnica para cada situação. Foram elas:

- **Support Vector Machine (SVM):** É uma técnica que busca traçar retas (ou curvas) entre as categorias num plano cartesiano (ZENG et al., 2008). Ou seja, busca delimitar o espaço de cada classe no gráfico. Dessa forma, cada dado novo pode ser colocado no gráfico e classificado de acordo com a limitação de espaço inicial.
- **K-Nearest Neighbors (KNN):** É uma técnica que calcula a distância de um novo dado para os dados já conhecidos (SHAHINFAR; KELMAN; KAHN, 2019). Assim, ele consegue achar os vizinhos mais próximos do dado não classificado e então classificá-lo como de mesma classe em função da relação de vizinhança.
- **Logistic Regression (LR):** A regressão logística é uma técnica que faz uma estimativa da probabilidade de novo dado pertencer a uma determinada classe, em detrimento de um conjunto de classes possíveis.
- **Linear Discriminant Analysis (LDA):** Baseado em operações estatísticas, trata-se de uma técnica que calcula a probabilidade de o novo dado pertencer às classes conhecidas, e assim retorna a classe com maior probabilidade resultante (ZHAO; WANG; NIE, 2018).
- **Classification and Regression Tree (CART):** É uma técnica que cria uma árvore binária que combina todas as características de entrada (ou features) para chegar a um resultado (ZHU et al., 2018). Dessa forma, se baseia em decisões de sim/não para chegar nos resultados.
- **Naive Bayes (NB):** De forma semelhante à técnica LDA, calcula a probabilidade de

um dado pertencer ou não a uma categoria, sendo amplamente utilizado em classificação e filtragem de textos.

- **Gradient Boosting Tree (GBT):** É uma técnica que fornece resultados interpretáveis, exigindo pouco pré-processamento de dados e ajuste dos parâmetros. Funciona como um algoritmo iterativo que combina funções parametrizadas simples com desempenho baixo para produzir uma regra de previsão altamente precisa (GUELMAN, 2012).

- **eXtreme Gradient Boosting (XBG):** É um eficiente solucionador de modelo linear e um algoritmo de aprendizado baseado em árvore. Oferece suporte a várias funções objetivas, incluindo regressão e classificação. Trata-se de uma solução desenvolvida para ser extensível, de modo que os usuários também possam definir e implementar seus próprios objetivos (CHEN et al., 2015).

- **Random Forest (RF):** É um método que constrói múltiplas árvores de decisão, de forma aleatória, a partir dos dados de treinamento. A decisão acerca de uma predição é tomada a partir da análise automatizada sobre as árvores que convergem para um mesmo valor ou classe. O erro de generalização converge para um limite conforme o número de árvores presentes aumenta (BREIMAN, 2001).

A aplicação dessas técnicas de predição teve como objetivo maximizar aspectos que são considerados importantes no momento de comercialização da carne animal. Para isto, foram selecionadas três características consideradas cruciais para o desenvolvimento esperado, com base no banco de dados constituído. Além disso, foi visto que diversos trabalhos da literatura realizam a predição de características relacionadas à carcaça animal (SHAHINFAR et al., 2019).

Esta pesquisa compreendeu, portanto, o desenvolvimento de três modelos de predição, utilizando AM, onde o objetivo de cada modelo foi predizer as seguintes características dos ovinos, com base nos dados históricos tabulados:

- Ganho Médio Diário (GMD);
- Consumo de Matéria Seca (CMS);
- Peso Corporal Final (PCF).

Para desenvolver os modelos de predição do GMD, CMS e do PCF como variáveis dependentes ou alvo (outputs) (Quadro 2), utilizou-se uma base de dados composta por 61 colunas, onde cada coluna representa uma variável. Vale ressaltar que cada linha representa uma média de tratamento, como explicado no início desta seção.

Dentre os 61 parâmetros disponíveis, foram selecionados os mais relevantes para cada característica do animal a ser predita. Para as predições do GMD foram selecionadas 3 variáveis independentes para as entradas (inputs) (Quadro 2) no desenvolvimento dos modelos, sendo elas: NDT dieta, PCi e PCf. Já para CMS, foram selecionadas GMD e PC metabólico como características para as entradas no desenvolvimento de modelos de AM. E, para a predição do PCf, foram utilizados apenas o PCi e os Dias em confinamento (DEC).

As características selecionadas foram escolhidas com base em sua capacidade de provocar uma resposta biológica no animal e em sua relevância para o objetivo da análise. Isso significa que foram identificadas as características biológicas que seriam mais adequadas para a análise em questão, tendo sido selecionadas dentre outras características possíveis.

A escolha dessas características foi baseada em uma avaliação criteriosa de sua capacidade de desencadear uma resposta biológica no animal em questão, ou seja, sua capacidade de fornecer informações relevantes sobre o processo biológico que se deseja investigar. Por exemplo, se o objetivo da análise é estudar o impacto de um medicamento no organismo animal, pode ser importante selecionar características que indiquem a absorção, distribuição e eliminação desse medicamento, bem como seus efeitos sobre a função celular ou a resposta imunológica.

Por outro lado, se o objetivo é avaliar o estresse do animal, características como os níveis de hormônios relacionados ao estresse e alterações comportamentais podem ser mais relevantes para a análise.

Em resumo, a seleção cuidadosa das características biológicas mais adequadas para uma análise é fundamental para garantir a obtenção de resultados precisos e relevantes.

Quadro 2. Tabela de variáveis utilizadas nos modelos de predição utilizando AM.

Variáveis independentes	Variáveis Dependentes (alvo)		
	Ganho Médio Diário (GMD)	Consumo de Matéria Seca (CMS)	Peso Corporal Final (PCf)
NDT dieta, PCi e PCf	GMD e PC metabólico	PCi e Dias em experimento	

GMD: Ganho Médio Diário (g/dia); **CMS:** Consumo de Matéria Seca (g/dia); **PCf:** Peso Corporal Final (kg); **NDT dieta:** Nutrientes Digestíveis Totais da Dieta (g/kg); **PCi:** Peso Corporal Inicial (kg); **PC metabólico:** Peso Corporal metabólico; **Dias em experimento:** Dias em experimento (dias).

3.4 Ferramentas utilizadas e desenvolvimento da RNA

O Google Colab¹ foi utilizado como ferramenta para implementação e execução dos algoritmos das técnicas selecionadas neste estudo, já que ele permite que sejam produzidos códigos na linguagem Python e estes sejam executados diretamente na infraestrutura do Google. Considerando sua praticidade e, também, por se tratar de uma ferramenta que está sendo muito utilizada em tarefas relacionadas à ciência de dados, o Google Colab foi adotado para as tarefas envolvendo análise da base de dados, desenvolvimento e avaliação dos modelos.

Primeiramente, foram criados arquivos no Colab para cada preditor a ser desenvolvido, de modo que cada um pudesse ser analisado separadamente; ressaltando-se que foi necessário familiarizar-se com os dados. Após importar a base de dados para o Colab, foi verificada a presença de valores nulos, ou seja, ausência de dados em alguma coluna. Neste estudo exploratório dos dados, apenas um valor nulo foi detectado, o qual foi substituído pela média dos valores da coluna, utilizando métodos da biblioteca Pandas². Essa forma de tratamento de valores nulos foi escolhida devido ao pequeno volume de dados, no qual excluir uma linha inteira representaria uma perda significativa de informações.

Seguindo a análise exploratória dos dados, utilizou-se a biblioteca Matplotlib³ com o intuito de entender visualmente como as informações estavam dispostas. Estas informações foram exibidas em histogramas e diagramas de caixa, nos quais foi possível observar cada aspecto isoladamente. Com isso, foi possível conhecer a distribuição das colunas, observar padrões e também analisar outliers; ou seja, pontos discrepantes nas amostras, que foram removidos quando não atenderam aos critérios, como forma de diminuir os ruídos. Além disso, a partir dessa análise, foi possível definir qual a melhor estratégia para realizar o pré-processamento de informações.

Com o objetivo de realizar as previsões do GMD, CMS e PCf a partir do registro histórico nutricional de vários ovinos, foram experimentadas dez técnicas clássicas de aprendizado de máquina: KNN, SVM, LR, LDA, CART, NB, GBT, XBG, RF e RNA. Os modelos de predição foram treinados para posteriores avaliações. O melhor modelo preditor para cada característica foi selecionado com base na performance calculada por meio do F1-score.

¹ <https://colab.research.google.com/>

² <https://pandas.pydata.org/>

³ <https://matplotlib.org/>

O F1-score, também encontrado na literatura como F1-measure, nada mais é do que uma medida de precisão de um teste. Ele representa a média harmônica entre a precisão e o recall, podendo seu valor ir de zero a um. Valores altos (próximos de 1,0 ou de 100%) indicam um alto desempenho de classificação (THARWAT, 2018). A acurácia indica uma performance geral do modelo criado, ou seja, se analisarmos todas as classificações que o modelo executou, e seu valor se baseará em quantas ele fez corretamente dentre todas as classificações. Já a precisão representa o resultado da quantidade de verdadeiros positivos sobre a soma de todos os valores positivos. Por último, o recall, também conhecido como sensibilidade ou revocação, representa a capacidade do método testado de detectar corretamente resultados classificados como positivos (MARIANO, 2021).

A avaliação da performance dos modelos de aprendizado de máquina baseados no F1-score é feita por meio da comparação dos valores obtidos por cada modelo treinado para a predição das características do gado. O F1-score é uma medida de precisão que leva em conta tanto a precisão quanto o recall, ou seja, a capacidade do modelo de prever corretamente tanto os verdadeiros positivos quanto os falsos negativos.

O F1-score varia de 0 a 1, sendo que um valor próximo de 1 indica um alto desempenho de classificação do modelo. Por exemplo, se um modelo tem um F1-score de 0,9 para a predição do GMD (ganho médio diário) de ovinos, isso significa que ele é capaz de prever corretamente 90% dos casos. Isso é um indicativo de que o modelo é capaz de fornecer uma previsão precisa e confiável para essa característica.

A escala do F1-score é bastante intuitiva. Um valor de 0 indica que o modelo não é capaz de prever corretamente nenhuma das características, enquanto um valor de 1 indica que o modelo é capaz de prever corretamente todas as características. Um valor intermediário indica um desempenho parcial do modelo.

Os valores do F1-score são obtidos a partir dos resultados do treinamento do modelo e da validação do modelo em um conjunto de dados de teste, que contém amostras que não foram usadas no treinamento do modelo. Para avaliar a performance do modelo, é necessário comparar os valores do F1-score obtidos para cada modelo treinado e selecionar o melhor modelo para cada característica com base nessa medida de desempenho.

Antes do desenvolvimento dos modelos de AM, foi realizado um pré-processamento adicional nos dados. Para realizar este pré-processamento, foi aplicado um algoritmo que refaz a escala sem alterar a natureza do dado. Essa técnica foi realizada para que o modelo trabalhasse com valores mais próximos possíveis de zero, sem perder suas características.

Além disso, evita que colunas com números maiores transmitam, erroneamente, maior importância, em detrimento de colunas com números relativamente pequenos. Após uma análise minuciosa, na qual foram testadas diferentes formas automatizadas de pré-processamento, foi observado que a melhor técnica foi a *StandardScaler*⁴, presente na biblioteca *preprocessing*⁵ do *scikit-learn*⁶, que refaz cada coluna e tenta remodelar para uma distribuição normal. Com isso, ao final, a predição dos modelos desenvolvidos terá um erro de aproximação numérica.

Além disso, para treinar a RNA, a variável alvo de cada preditor sofreu ainda outro tipo de pré-processamento: o uso da técnica de *One Hot Encoding*⁷, também disponível na biblioteca *preprocessing*, citada anteriormente. Essa técnica consiste em transformar um conjunto de valores categóricos em uma matriz de valores binários. Como por exemplo: para o conjunto {1,2,3,4}, após aplicar *One Hot Encoding*⁷, a sequência (0,0,0,1) representaria o número 4, a sequência (0,1,0,0) representaria o número 2, e assim por diante, codificando todo o conjunto de rótulos. Esta técnica foi necessária, pois a implementação da RNA retorna uma saída com um vetor binário de probabilidades, ou seja, caso a saída retornada pelo preditor seja o elemento (ou rótulo) 2 com 80% de certeza, poderia ter como saída algo como: (0.05, 0.80, 0.06, 0.09).

Em seguida, foi realizado o treinamento e avaliação dos preditores com base nas técnicas de AM abordadas neste estudo. Para isto, foi utilizada a técnica de Validação Cruzada Estratificada (VCE) com quatro dobras, também importada do *scikit-learn*, que consiste em executar o mesmo algoritmo quatro vezes (4-folds), com subconjuntos diferentes de dados, e considerar a média da performance observada nas quatro execuções. Devido ao baixo volume de dados disponíveis para treino e teste neste estudo, a utilização da VCE com quatro dobras foi a melhor opção para garantir uma avaliação robusta do desempenho dos modelos. O valor quatro foi escolhido, pois é o máximo permitido na função programática que realiza a VCE, devido ao baixo volume de dados disponíveis para treino e teste. Dessa forma, a VCE evita que ocorram casos em que o algoritmo tenha um desempenho muito bom devido a aleatoriedade dos dados, ou ainda um desempenho muito ruim pelo mesmo motivo. A Figura 1 descreve um exemplo de VCE com 5 dobras. É possível notar que, apesar de divididos, cada bloco de dados mantém a proporcionalidade do conjunto como um todo, com relação a suas

⁴ <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html>

⁵ <https://scikit-learn.org/stable/modules/classes.html#module-sklearn.preprocessing>

⁶ <https://scikit-learn.org/stable/index.html>

⁷ <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.OneHotEncoder.html>

classes.

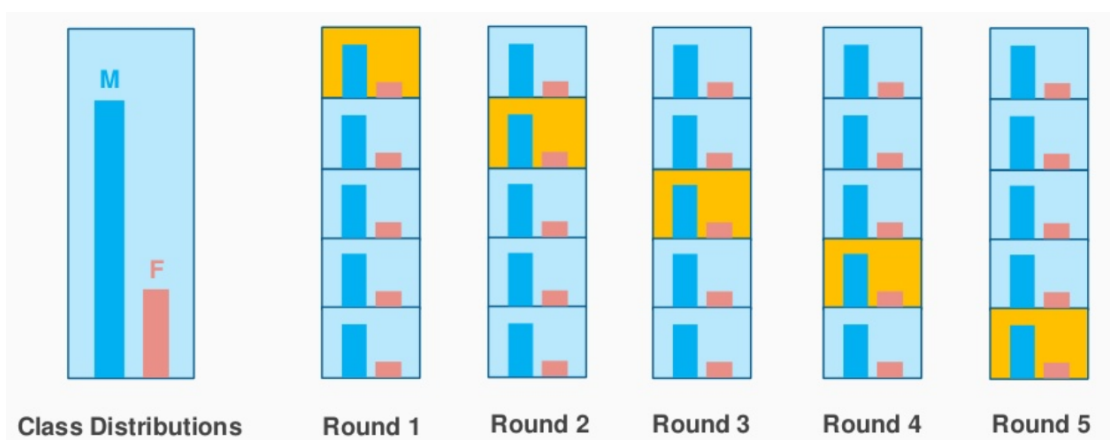


Figura 1. Exemplo de validação cruzada estratificada 5-fold.
Fonte: Stack Exchange⁸, 2013.

Após o treinamento, os modelos foram avaliados utilizando os dados de teste, que foram separados dos dados de treino. A métrica de desempenho utilizada foi baseada em acurácia, precisão, recall e F1-score, que são métricas amplamente utilizadas em aprendizado de máquina para avaliar a qualidade dos modelos. Utilizando também o scikit-learn, as métricas de acurácia, precisão, recall e F1-score foram calculadas em cada execução de um modelo e armazenadas para comparação.

Ao final dessa etapa, foram produzidos, no total, 120 modelos, pois ao utilizar a VCE com 4 dobras, foi possível gerar 4 modelos para cada uma das 10 técnicas exploradas, em cada uma das 3 variáveis a serem preditas.

4. RESULTADOS E DISCUSSÃO

Neste estudo, logo após concluir as etapas de treino e teste dos 120 modelos desenvolvidos, sendo quarenta para cada um dos três preditores, foi possível visualizar e analisar os resultados preliminares obtidos.

Em relação aos inputs utilizados para obtenção dos resultados da saída, o estudo em questão utilizou uma abordagem de seleção de variáveis para identificar as mais relevantes para cada característica do animal a ser predita. Dessa forma, foram selecionadas as variáveis mais importantes para cada modelo de predição, sendo que para o GMD foram utilizadas as variáveis NDT dieta, PC_i e PC_f, para o CMS foram utilizadas as variáveis GMD e PC

⁸ <https://stats.stackexchange.com/questions/49540/understanding-stratified-cross-validation>

metabólico, e para a predição do PCf, foram utilizadas apenas o PCi e os Dias em confinamento (DEC). É importante ressaltar que as variáveis nutricionais foram sim utilizadas, sendo que as variáveis NDT dieta e PC metabólico são exemplos de variáveis nutricionais utilizadas no modelo. O processo de seleção das variáveis a serem utilizadas é essencial para a construção de modelos de predição precisos, pois nem todas as variáveis disponíveis são relevantes para a predição da característica desejada.

A Figura 2 apresenta um gráfico comparando a performance média das técnicas de IA exploradas para predizer o GMD, uma vez que cada uma passou por uma VCE com quatro dobras. Logo, para predizer o GMD, foram quatro modelos gerados e avaliados para cada técnica explorada. O F1-score foi escolhido para comparar performance pois, por representar uma média harmônica entre Precisão e Recall, no qual consegue apresentar uma informação balanceada que auxilia na comparação entre o desempenho dos modelos. Nesse modelo, a RNA apresentou o menor desempenho em comparação aos outros algoritmos. Os algoritmos LDA, SVM, XBG e RF apresentaram a melhor resposta em relação à acurácia, no entanto, o GBT apresentou o melhor F1-score (acima de 0,4) em relação aos demais algoritmos, apesar dos demais algoritmos estarem abaixo de 0.5.

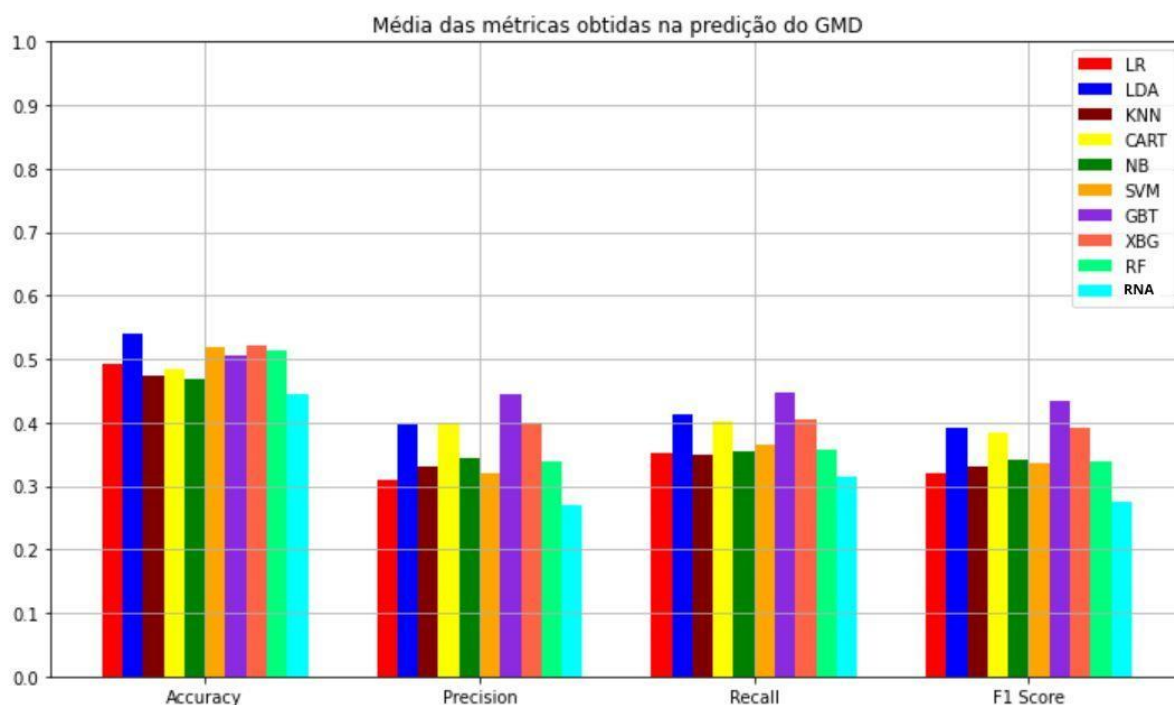


Figura 2. Gráfico de barras referentes à acurácia, precisão, recall e ao F1-score das técnicas de predição do GMD.

O gráfico de comparação de performance média dos modelos de predição do consumo

de matéria seca (CMS) é apresentado na Figura 3. Observa-se que a predição do CMS teve melhor acurácia (acima de 0,5) representada pelos algoritmos LDA, KNN, CART, NB, SVM, XGB e RF; contudo, o CART, SVM, KNN e RF para as demais métricas (Precision, Recall e F1 Score). No entanto, a RNA apresentou o menor (variando entre 0,2 a 0,4) desempenho em todas as métricas avaliadas.

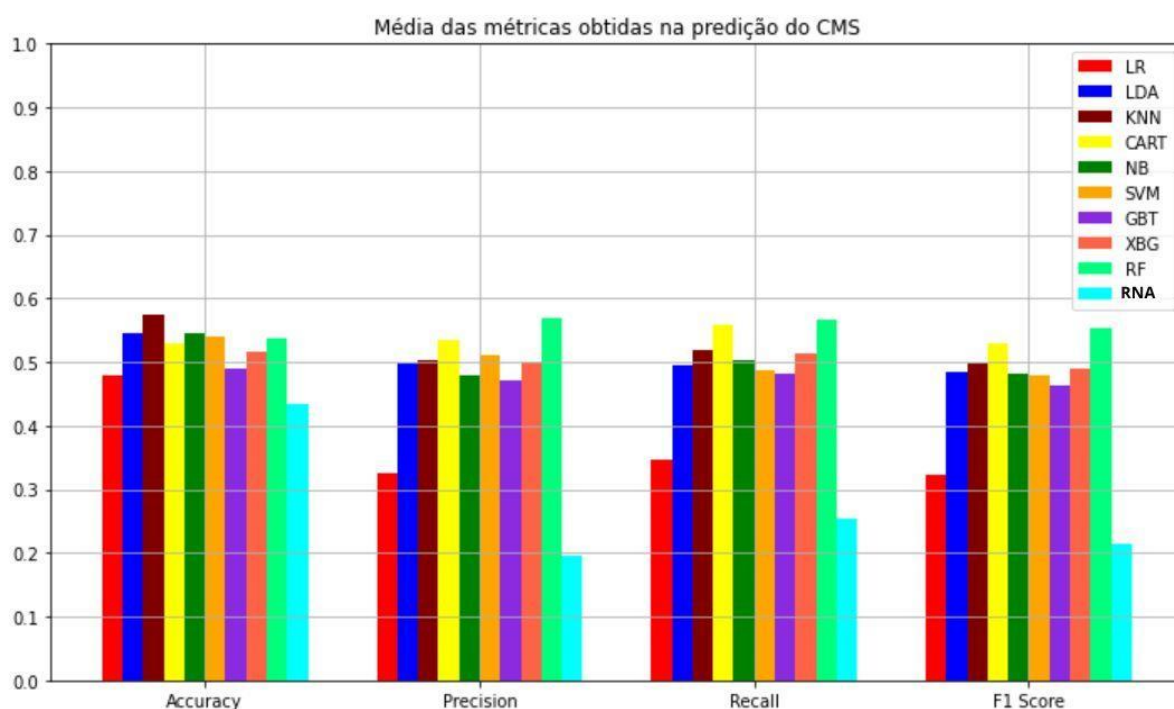


Figura 3. Gráfico de barras referentes à acurácia, precisão, recall e ao F1 Score das técnicas de predição do CMS.

Na Figura 4 estão apresentadas as métricas do modelo para PCf, em que também observou-se que a RNA apresentou desempenho muito abaixo (variando de 0,2 a 0,4) do esperado para todas as métricas estudadas, embora não tenha sido a pior técnica em termos de acurácia. Por outro lado, os algoritmos GBT, XGB e RF apresentaram os melhores desempenhos para acurácia, precisão, recall e F1-score.

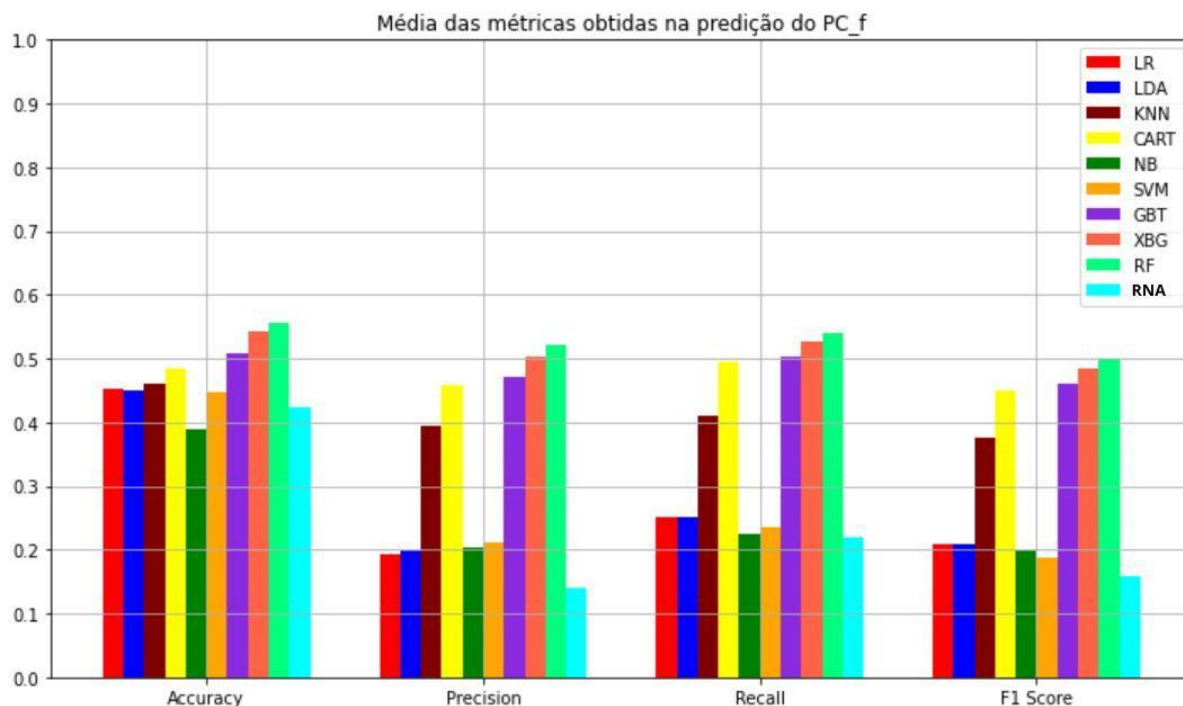


Figura 4. Gráfico de barras referentes à acurácia, precisão, recall e ao F1 Score das técnicas de predição do PCf.

Para todos os pontos preditos, GMD, CMS e PCf, o algoritmo da RNA apresentou um desempenho menor quando comparado a outros algoritmos incluídos nas análises. No entanto, isso pode ser uma consequência do baixo número de dados disponíveis para realização do treinamento, verificação e checagem dessa ferramenta.

Segundo Kominakis et al. (2002), um maior conjunto de dados na fase de treinamento da RNA poderia melhorar a qualidade da predição da rede, pois possibilitaria direcionar melhor os parâmetros específicos, ou seja, quanto maior o número de dados disponíveis para o treinamento do algoritmo, maior será sua habilidade de aprendizagem para predição; isso, provavelmente melhoraria o desempenho observado da RNA em relação aos demais algoritmos analisados.

Pereira e Centeno (2017), avaliando o tamanho de amostra de treinamento para RNA na classificação supervisionada de imagens utilizando dados espectrais e laser scanner, observaram que a medida que aumentavam o número de amostras de treinamento da RNA, ocorria um aumento dos acertos globais do algoritmo. Os autores justificaram que ao aumentar a quantidade de elementos na amostra de treinamento, aumenta-se também a variabilidade dos valores fornecidos a RNA, o que acaba aumentando a capacidade do algoritmo de discriminar pequenas variações em valores.

No entanto, os autores afirmam que embora haja uma tendência de crescimento dos

acertos, essa não necessariamente será linear e constante com o aumento das amostras utilizadas para treinar a rede. Em contrapartida, o que pode ocorrer é que à medida que se aumenta a variabilidade de informações dos dados apresentados à RNA, melhora-se a capacidade de discriminar novos elementos. Logo, não basta apenas possuir mais dados, eles também precisam de qualidade, visando reduzir esse ruído e, conseqüentemente, os efeitos negativos que podem trazer no aprendizado de máquina.

O desempenho esperado da RNA depende da natureza dos dados e do contexto da aplicação. Em alguns casos, as RNAs podem ter um desempenho superior em relação a outras técnicas de aprendizado de máquina, especialmente em problemas complexos e não lineares. No entanto, em outros casos, pode haver limitações relacionadas à quantidade e qualidade dos dados disponíveis, além da escolha adequada da arquitetura e dos parâmetros da rede.

Um estudo realizado por Oliveira et al. (2019) avaliou o desempenho de diferentes técnicas de aprendizado de máquina, incluindo RNAs, para a predição de características de ovinos com base em dados nutricionais. Os autores encontraram que as RNAs tiveram um desempenho inferior em relação a outras técnicas, como Random Forest e Gradient Boosting Trees, para a predição de algumas características, como ganho médio diário e consumo de matéria seca. Os resultados indicaram que as RNAs não conseguiram capturar a complexidade dos dados nutricionais dos ovinos de forma eficiente.

Em outro estudo, realizado por Zambrano et al. (2017), as RNAs foram utilizadas para a predição do ganho de peso de suínos com base em dados de nutrição e genética. Os autores encontraram que as RNAs tiveram um desempenho superior em relação a outras técnicas, como SVM e Random Forest, para a predição do ganho de peso em suínos. No entanto, os autores destacaram que a performance das RNAs pode ser limitada pela quantidade e qualidade dos dados disponíveis, especialmente quando há poucas amostras de treinamento.

Outro ponto que pode ter colaborado para o baixo desempenho de predição da RNA possivelmente foi a ausência de informações no banco de dados utilizado. Alguns dos materiais usados para compor o banco de dados não apresentavam todas as informações completas, mas foram mantidos a fim de evitar redução e remoção da linha, mesmo aqueles trabalhos em que algumas informações não puderam ser estimadas utilizando fórmulas específicas. Assim, as ausências foram superadas com inserção de médias dos valores existentes em cada coluna. Além disso, umas das medidas adotadas para que uma coluna fosse analisada, era que a coluna tivesse mais de 75% de dados preenchidos, pois essas ausências criadas, possivelmente poderiam causar incerteza no processo de aprendizagem

(BETKER; SZTURM; MOUSSAVI, 2003).

Assim, uma possível melhoria no desempenho seria a partir da expansão, com a busca por mais trabalhos, não se restringindo apenas aos realizados no Departamento de Zootecnia da UFRPE, mas sim de trabalhos disponíveis na literatura, e conseqüentemente, aumentando o banco de dados para ser utilizado no treinamento, aprendizagem e validação da Rede Neural.

No entanto, outro ponto que pode ser investigado em trabalhos futuros é o conjunto de variáveis que foram selecionadas para aplicação das técnicas. Como foram selecionadas características sob o ponto de vista de gerar resposta biológica no animal, estas variáveis podem não ter sido o conjunto mais eficiente em relação à parte computacional, podendo ter causado confusão no modelo, pois mesmo outras técnicas se apresentando superiores ao desempenho observado da RNA, essas ainda obtiveram um desempenho muito baixo, com F1-Score inferior a 0,5 na maioria dos casos.

Com isso, três outras alternativas podem ser colocadas para futuros trabalhos. A primeira delas, seria manter o conjunto de características selecionadas para predição e expandir o banco de dados com mais exemplos/registros, buscando trabalhos que tragam, de preferência, a média dos animais e não a média de um conjunto de animais, permitindo a aplicação das técnicas com um conjunto maior de dados. A segunda alternativa seria manter o banco de dados atual e estudar outras características que poderiam ser utilizadas para predição do GMD, CMS e PCf. Por fim, a terceira alternativa seria o melhor cenário de investigação, que consiste em buscar mais registros de dados para expandir o banco de dados atual, além de realizar uma análise para selecionar melhores características a serem utilizadas no desenvolvimento dos modelos de predição, permitindo uma análise comparativa do desempenho dos modelos em relação às características biológicas.

5. CONCLUSÃO

Este estudo avaliou o desempenho de diferentes técnicas de aprendizado de máquina, incluindo Redes Neurais Artificiais (RNA), para a predição de características de ovinos com base em dados nutricionais. Os resultados mostraram que as RNAs apresentaram desempenho abaixo do esperado, assim como as demais técnicas testadas. No entanto, como este é um estudo preliminar, alguns entraves foram detectados e podem ser corrigidos em trabalhos futuros.

Novas pesquisas precisam ser realizadas para explorar melhor as possibilidades das

RNAs na área da ciência animal, levando em conta o número e a qualidade dos dados disponíveis e o conjunto de variáveis selecionado para predição.

Uma possível alternativa para as RNAs, quando se trata de um pequeno volume de dados, é o uso de outras técnicas como o Random Forest (RF), que apresentou melhores resultados na maior parte das análises neste estudo. No entanto, é importante que futuras pesquisas sejam realizadas para verificar e validar esses resultados.

Em resumo, embora as RNAs possam ser uma técnica promissora na área da ciência animal, seus resultados podem ser limitados pela quantidade e qualidade dos dados disponíveis. Por isso, é importante avaliar cuidadosamente a seleção de variáveis e técnicas adequadas para cada caso específico, visando obter resultados mais precisos e confiáveis.

6. REFERÊNCIAS

- AGUIAR, S. R. **Substituição do milho e do farelo de soja por levedura e uréia na alimentação de ovinos**. 2007. Dissertação (Mestrado em Zootecnia) - Universidade Federal Rural de Pernambuco, Departamento de Zootecnia, Recife.
- ALCANTARA, L. M. et al. Machine learning classification of breeding protocol descriptions from Canadian Holsteins. **Journal of Dairy Science**, v. 105, ed. 10, p. 8177–8188, 2022. DOI <https://doi.org/10.3168/jds.2021-21663>.
- ALVES, F. J. L. **Fontes proteicas alternativas ao farelo de soja na alimentação de ovinos**. 2012. Dissertação (Mestrado em Zootecnia) - Universidade Federal Rural de Pernambuco, Departamento de Zootecnia, Recife.
- ANDRADE, R. B. **Silagem de cana-de-açúcar (*Saccharum ssp.*) aditivada com resíduo do beneficiamento do feijão (*Phaseolus vulgaris L.*) na dieta de ovinos em confinamento**. 2010. Dissertação (Mestrado em Zootecnia) - Universidade Federal Rural de Pernambuco, Departamento de Zootecnia, Recife.
- ANDRADE, S. F. J. **Palma forrageira (*Nopalea cochenillifera Salm-Dyck*) in natura ou farelada na dieta de borregos**. 2010. Dissertação (Mestrado em Zootecnia) - Universidade Federal Rural de Pernambuco, Departamento de Zootecnia, Recife.
- ANDRADE, R. P. X. **Casca de mamona em substituição ao feno de capim tifton: consumo, digestibilidade e desempenho de ovinos**. 2011. Dissertação (Mestrado em Zootecnia) - Universidade Federal Rural de Pernambuco, Departamento de Zootecnia, Recife.
- BEZEN, R.; EDAN, Y.; HALACHMI, I. Computer vision system for measuring individual cow feed intake using RGB-D camera and deep learning algorithms. **Computers and Electronics in Agriculture**, v. 172, 2020. DOI <https://doi.org/10.1016/j.compag.2020.105345>.
- BETKER, A. L.; SZTURM, T.; MOUSSAVI, Z. Application of feedforward backpropagation neural network to center of mass estimation for use in a clinical environment. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No.03CH37439), 25, 2003, Cancun, Mexico. **Proceedings**. [s.l.]: IEEE, 2003, v. 3, p. 2714-2717. DOI <https://doi.org/10.1109/IEMBS.2003.1280477>.
- BEZERRA, S. B. L. **Inclusão da palma forrageira (*Nopalea cochenillifera (L.) Salm-Dyck*) em dietas de ovinos em crescimento**. 2015. Tese (Doutorado em Zootecnia) - Universidade Federal Rural de Pernambuco, Departamento de Zootecnia, Recife.
- BEZERRA, W. A. O. et al. Utilização de redes neurais artificiais na predição de estresse em ovelhas prenhas. In: Encontro Internacional de Gestão, Desenvolvimento e Inovação (EIGEDIN), 4, 2020, online. **Anais**. [s.l.]: EIGEDIN, 2020, v. 4, ed. 1. Disponível em: <https://periodicos.ufms.br/index.php/EIGEDIN/article/view/11619>. Acesso em: 10 nov. 2022.
- BREIMAN, L. Random Forests. **Machine Learning**, v. 45, p. 5–32, 2001. DOI

<https://doi.org/10.1023/A:1010933404324>.

CARDOSO, D. B. **Caracterização da carcaça e da carne de cordeiros alimentados com borra de manipueira em substituição ao milho**. 2014. Dissertação (Mestrado em Zootecnia) - Universidade Federal Rural de Pernambuco, Departamento de Zootecnia, Recife.

CARDOSO, D. B. **Inclusão de palma miúda (*Nopalea cochenilifera salm dyck*) na dieta de cordeiros**. 2018. Tese (Doutorado Integrado em Zootecnia) - Universidade Federal Rural de Pernambuco, Universidade Federal da Paraíba, Universidade Federal do Ceará, Departamento de Zootecnia da UFRPE, Recife.

CHEN, T. et al. Xgboost: extreme gradient boosting. **R package version 0.4–2**, 2015.

COELHO, M. C. S. C. **Feno de pornunça (*Manihot spp.*) na alimentação de ovinos em confinamento no semiárido**. 2014. Tese (Doutorado Integrado em Zootecnia) - Universidade Federal Rural de Pernambuco, Universidade Federal da Paraíba, Universidade Federal do Ceará, Departamento de Zootecnia da UFRPE, Recife.

COSTA, E. J. X.. Inteligência artificial aplicada à Zootecnia. **Revista Brasileira de Zootecnia**, v. 38, p. 390-396, 2009. DOI <https://doi.org/10.1590/S1516-35982009001300038>.

CUNHA, M. G. G. **Aspectos nutricionais, produtivos e reprodutivos em ovinos alimentados com rações contendo níveis crescentes de caroço de algodão**. 2006. Tese (Doutorado em Zootecnia) - Universidade Federal Rural de Pernambuco, Departamento de Zootecnia, Recife.

FERNANDES, S. R. et al. Características de desempenho e eficiência alimentar de touros Purunã em crescimento de três classes de consumo alimentar residual. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v. 66, ed. 1, p. 268-276, 2014. Disponível em: <https://www.scielo.br/j/abmvz/a/CTsQfHQ3gXvzXBhDmznKKmj/?lang=pt&format=pdf.v> Acesso em: 1 nov. 2022.

FERRAZ, L. V. **Fontes de proteína em dietas à base de palma forrageira para ovinos em terminação**. 2013. Dissertação (Mestrado em Zootecnia) - Universidade Federal Rural de Pernambuco, Departamento de Zootecnia, Recife.

FREITAS, A. P. D. **Mazoferm**: alternativa ao farelo de soja na dieta de ovinos. 2013. Dissertação (Mestrado em Zootecnia) - Universidade Federal Rural de Pernambuco, Departamento de Zootecnia, Recife.

GUELMAN, L. Gradient boosting trees for auto insurance loss cost modeling and prediction. **Expert Systems with Applications**, v. 39, ed. 3, p. 3659-3667, 2012. DOI <https://doi.org/10.1016/j.eswa.2011.09.058>.

HERBSTER, C. J. L. et al. Weight adjustment equation for hair sheep raised in warm conditions. **Animal**, v. 14, ed. 8, p. 1718-1723, 2020. DOI <https://doi.org/10.1017/S1751731120000294>.

KOMINAKIS, A. P. et al. A preliminary study of the application of artificial neural networks

to prediction of milk yield in dairy sheep. **Computers and Electronics in Agriculture**, v. 35, ed. 1, p. 35-48, 2002. DOI [https://doi.org/10.1016/S0168-1699\(02\)00051-0](https://doi.org/10.1016/S0168-1699(02)00051-0).

LACROIX, R. et al. Prediction of cow performance with a connectionist model Trans. **ASAE**, 38, 5, pp. 1573-1579, 1995.

LIMA, A. M. C.; BEZERRA, L. R.; NASCIMENTO, C. N. B.; VIANA, J. H. M.; LIRA, M. A.; FERNANDES, H. J. Comportamento de ovinos Santa Inês alimentados com dietas contendo diferentes níveis de fibra em detergente neutro. **Revista Brasileira de Saúde e Produção Animal**, v. 15, ed. 3, p. 589-601, 2014.

LIU, X. et al. A review of artificial neural networks in the constitutive modeling of composite materials. **Composites Part B: Engineering**, v. 224, p. 109-152, 2021. DOI <https://doi.org/10.1016/j.compositesb.2021.109152>.

MACIEL, M. V. **Utilização de feno ou silagem de maniçoba em substituição ao feno de Tifton-85 na alimentação de ovinos**. 2012. Dissertação (Mestrado em Zootecnia) - Departamento de Zootecnia, Universidade Federal Rural de Pernambuco, Recife.

MAIOR JÚNIOR, R. J. S. **Substituição de feno de tifton pela cana-de-açúcar sobre o rendimento, características de carcaça e componentes não carcaça em ovinos em confinamento**. 2008. Dissertação (Mestrado em Zootecnia) - Universidade Federal Rural de Pernambuco, Departamento de Zootecnia, Recife.

MARIANO, D. Métricas de avaliação em machine learning: acurácia, sensibilidade, precisão, especificidade e F-score. **BIOINFO - Revista Brasileira de Bioinformática e Biologia Computacional**, v. 15, ed. 1, 2021. DOI <http://dx.doi.org/10.51780/978-6-599-275326-15>. Disponível em: https://www.researchgate.net/publication/353001655_Metricas_de_avaliacao_em_machine_learning_acuracia_sensibilidade_precisao_especificidade_e_F-score. Acesso em: 8 fev. 2023.

MATTOS, C. W. **Associação de palma forrageira (*Opuntia ficus-indica* Mill) e feno de erva-sal (*Atriplex nummularia* L) em dietas para cordeiros Santa Inês em confinamento**. 2009. Tese (Doutorado em Zootecnia) - Universidade Federal Rural de Pernambuco, Departamento de Zootecnia, Recife.

MORAIS, J. S. **Mistura do resíduo da indústria de doce e farelo de glúten de milho na alimentação de cordeiros em crescimento**. 2020. Tese (Doutorado em Zootecnia) - Universidade Federal Rural de Pernambuco, Recife.

MOURA, M. S. C. **Feno de maniçoba (*Manihot pseudoglaziovii* Muel Arg.) e palma forrageira (*Nopalea cochenillifera* Salm Dick), na dieta de ovinos em crescimento**. 2013. Tese (Doutorado Integrado em Zootecnia) - Universidade Federal Rural de Pernambuco, Universidade Federal da Paraíba, Universidade Federal do Ceará, Departamento de Zootecnia da UFRPE, Recife.

OLIVEIRA, A. B. **Substituição da silagem de sorgo por resíduo da cultura de milho na alimentação de cordeiros confinados**. 2018. Tese (Doutorado em Zootecnia) - Universidade Federal Rural de Pernambuco, Programa de Pós-Graduação em Zootecnia, Recife.

OLIVEIRA, A. P. et al. Meta-analysis of the energy and protein requirements of hair sheep raised in the tropical region of Brazil. **Journal of Animal Physiology and Animal Nutrition**, n. 0, p. 1–9, 2017. DOI <http://doi.wiley.com/10.1111/jpn.12700>.

OLIVEIRA, J. P. F. **Características de carcaça e qualidade da carne de ovinos da raça Santa Inês alimentados com mazoferm em substituição ao farelo de soja**. 2013. Dissertação (Mestrado em Zootecnia) - Universidade Federal Rural de Pernambuco, Departamento de Zootecnia, Recife.

OLIVEIRA, L. S.; SILVA, L. F.; BEZERRA, L. F.; COSTA, R. G.; ALBUQUERQUE, F. H. Comparison of machine learning algorithms for predicting performance and carcass characteristics of feedlot-finished lambs. **Computers and Electronics in Agriculture**, v. 156, p. 159-165, 2019.

OLIVEIRA, M. V. M. et al. Rendimento de carcaça, mensurações e peso de cortes comerciais de cordeiros Santa Inês e Bergamácia alimentados com dejetos de suínos em confinamento. **Revista Brasileira de Zootecnia**, v. 31, ed. 3, p. 1451-1458, 2002. DOI <https://doi.org/10.1590/S1516-35982002000600017>.

PANDORFI, H. et al. Uso de redes neurais artificiais para predição de índices zootécnicos nas fases de gestação e maternidade na suinocultura. **Revista Brasileira de Zootecnia**, v. 40, ed. 3, p. 676-681, 2011.

PEREIRA, C. C. O. **Principais métodos de avaliação da carcaça ovina**: revisão de literatura. 2016. Trabalho de Conclusão do Curso Zootecnia (Bacharelado em Zootecnia) - Escola de Medicina Veterinária e Zootecnia da Bahia, Universidade Federal da Bahia, Salvador.

PEREIRA, G. H. A.; CENTENO, J. A. S. Avaliação do tamanho de amostras de treinamento para redes neurais artificiais na classificação supervisionada de imagens utilizando dados espectrais e laser scanner. **Boletim de Ciências Geodésicas**, Curitiba, v. 23, ed. 2, p. 268-283, 2017. DOI <https://doi.org/10.1590/S1982-21702017000200017>.

SAMPERIO, E. et al. Lambs' live weight estimation using 3D images. **Animal: The international journal of animal biosciences**, v. 15, 2021. DOI <https://doi.org/10.1016/j.animal.2021.100212>.

SANTOS, K. C. **Efeito de diferentes fontes de carboidratos associadas à ureia em dietas para ruminantes**. 2017. Tese (Doutorado em Zootecnia) - Universidade Federal Rural de Pernambuco, Programa de Pós-Graduação em Zootecnia, Recife.

SANTOS FILHO, H. B. **Avaliação da manipueira em substituição ao milho na dieta de ovinos**. 2012. Dissertação (Mestrado em Zootecnia) - Universidade Federal Rural de Pernambuco, Departamento de Zootecnia, Recife.

SANT'ANA, D. A. et al. Weighing live sheep using computer vision techniques and regression machine learning. **Machine Learning with Applications**, v. 5, 2021. DOI <https://doi.org/10.1016/j.mlwa.2021.100076>.

SHADPOUR, S. et al. Predicting dry matter intake in Canadian Holstein dairy cattle using milk mid-infrared reflectance spectroscopy and other commonly available predictors via artificial neural networks. **Journal of Dairy Science**, v. 105, ed. 10, p. 8257–8271, 2022. DOI <https://doi.org/10.3168/jds.2021-21297>.

SHAHINFAR, S.; KELMAN, K.; KAHN, L. Prediction of sheep carcass traits from early-life records using machine learning. **Computers and Electronics in Agriculture**, v. 156, p. 159-177, 2019. DOI <https://doi.org/10.1016/j.compag.2018.11.021>.

SILVA, D. C. G. da. Predição de estresse em ovelhas prenhas e lactantes com o uso de redes neurais artificiais. **Ensaios e Ciência**, v. 25, ed. 2, p. 160-165, 2021. DOI <https://doi.org/10.17921/1415-6938.2021v25n2p160-165>. Disponível em: <https://ensaioseciencia.pgsskroton.com.br/article/view/8556>. Acesso em: 9 abr. 2023.

SILVA, E. C. **Substituição do milho por gérmen integral de milho na dieta de ovinos em confinamento**. 2012. Dissertação (Mestrado em Zootecnia) - Universidade Federal Rural de Pernambuco, Departamento de Zootecnia, Recife.

SILVA, F. L. R.; SANTANA, M. A.; OLIVEIRA, A. S.; OLIVEIRA, L. G.; MORAIS, J. A. S.; MEDEIROS, G. R. Perfil sanitário e reprodutivo de ovelhas Santa Inês criadas no sertão da Paraíba. **Revista Científica de Medicina Veterinária**, v. 16, ed. 1, p. 32-39, 2018.

SILVA, N. V. **Utilização do subproduto agroindustrial da goiaba (*Psidium Guajava L.*) na alimentação de ovinos em confinamento**. 2012. Tese (Doutorado em Zootecnia) - Centro de Ciências Agrárias, Universidade Federal da Paraíba, Areia.

SILVA, J. R. C. **Feno de alfafa em dietas para ovinos em crescimento**. 2015. Tese (Doutorado Integrado em Zootecnia) - Universidade Federal Rural de Pernambuco, Departamento de Zootecnia, Recife.

SILVA, F. J. S. **Componentes do peso corporal e qualidade da carne de ovinos Santa Inês alimentados com dietas contendo diferentes fontes de carboidratos associadas à ureia**. 2017. Dissertação (Mestrado em Zootecnia) - Universidade Federal Rural de Pernambuco, Programa de Pós-Graduação em Zootecnia, Recife.

SOUZA JÚNIOR, E. L. **Tamanho da estrutura corporal e o desempenho reprodutivo e produtivo de ovinos da raça Santa Inês**. 2007. Tese (Doutorado em Zootecnia) - Centro de Ciências Agrárias, Universidade Federal da Paraíba, Areia.

SOUSA, R. V. et al. Predictive model based on artificial neural network for assessing beef cattle thermal stress using weather and physiological variables. **Computers and Electronics in Agriculture**, Volume 144, Pages 37-43, 2018. DOI <https://doi.org/10.1016/j.compag.2017.11.033>.

THARWAT, A. Classification assessment methods. **Applied Computing and Informatics**, v. 17, ed. 1, p. 168-192, 2018. DOI <https://doi.org/10.1016/j.aci.2018.08.003>. Disponível em: <https://www.emerald.com/insight/content/doi/10.1016/j.aci.2018.08.003/full/html>. Acesso em: 8 fev. 2023.

URBANO, S. A. **Características de carcaça e qualidade da carne de ovinos alimentados com casca de mamona em substituição ao feno tifton**. 2011. Dissertação (Mestrado em Zootecnia) - Universidade Federal Rural de Pernambuco, Departamento de Zootecnia, Recife.

URBANO, S. A. **Fontes alternativas de energia para cordeiros em terminação**. 2015. Tese (Doutorado Integrado em Zootecnia) - Universidade Federal Rural de Pernambuco, Universidade Federal da Paraíba, Universidade Federal do Ceará, Departamento de Zootecnia da UFRPE, Recife.

VASCONCELOS, R. M. J. **Desempenho de ovinos alimentados com ração contendo diferentes níveis de cana-de-açúcar (sacharum ssp.) em substituição ao feno de tifton (cynodon dactylon)**. 2008. Dissertação (Mestrado em Zootecnia) - Universidade Federal Rural de Pernambuco, Departamento de Zootecnia, Recife.

VASCONCELOS, G. A. **Borra de manipueira em substituição ao milho na dieta de cordeiros**. 2013. Dissertação (Programa de Pós-Graduação em Zootecnia) - Universidade Federal Rural de Pernambuco, Departamento de Zootecnia, Recife.

XAVIER, T. C. **Características de carcaça de ovinos alimentados com fontes alternativas de proteína ao farelo de soja**. 2012. Dissertação (Mestrado em Produção Animal) - Universidade Federal Rural de Pernambuco, Departamento de Zootecnia, Recife.

XENOFONTE, A. R. B. **Desempenho e características de carcaça em ovinos mestiços em crescimento alimentados com diferentes níveis de farelo de babaçu (Orbiginea speciosa, Barb-RoOdr)**. 2006. Dissertação (Mestrado em Zootecnia) - Universidade Federal Rural de Pernambuco, Departamento de Zootecnia, Recife.

XU, A. et al. Applying artificial neural networks (ANNs) to solve solid waste-related issues: A critical review. **Waste Management**, v. 124, p. 385-402, 2021. DOI <https://doi.org/10.1016/j.wasman.2021.02.029>.

ZAMBRANO, L. I. P.; NÚÑEZ, R. A. S.; SARMIENTO, A. L. V.; BALCÁZAR, J. L. G. Comparison of machine learning algorithms to predict swine body weight. **Spanish Journal of Agricultural Research**, v. 15, ed. 3, 2017.

ZENG, Z. Q. et al. Fast training support vector machines using parallel sequential minimal optimization. In: International Conference on Intelligent System and Knowledge Engineering, 3, 2008, Xiamen, China. **Proceedings**. [s.l.]: IEEE, 2008. DOI <https://doi.org/10.1109/ISKE.2008.4731075>.

ZHAO, H.; WANG, Z.; NIE, F. A new formulation of linear discriminant analysis for robust dimensionality reduction. **Transactions on Knowledge and Data Engineering**, 2018, v. 31, p. 629–640. DOI <https://doi.org/10.1109/TKDE.2018.2842023>.

ZHU, F. et al. A classification algorithm of CART decision tree based on MapReduce attribute weights. **International Journal of Performability Engineering**, v. 14, p. 17-25, 2018. DOI <https://doi.org/10.23940/ijpe.18.01.p3.1725>.

ZURADA, J. M. Introduction to Artificial Neural Systems, **PWS Publishing Company**,

1992.