



MODELOS DE REDES NEURAIS PARA CLASSIFICAÇÃO DE COMPORTAMENTO E CONDIÇÃO DE SAÚDE DE BOVINOS

Palavras-Chave: Pecuária de Precisão, Classificação de dados, Redes Neurais.

Autores(as):

Laura Beatriz Costa da Silva, FT – UNICAMP

Prof. Dr. João Roberto Bertini Junior (orientador), FT - UNICAMP

INTRODUÇÃO:

A detecção precoce de qualquer problema de saúde é de grande importância para melhorar o bem-estar animal na pecuária intensiva. Muitas vezes, mesmo com os pecuaristas observando para detectar problemas de saúde e bem-estar, quando os sintomas se tornam evidentes, pode ser tarde demais para intervir, facilitando a propagação da doença [Ritter *et al.*, 2017].

Recentes avanços tecnológicos, têm permitido a captura e análise de sinais biológicos de animais, tornando possível identificar processos fisiológicos e comportamentais [Farooq *et al.*, 2022]. Tais possibilidades aumentaram o interesse de pecuaristas em novas soluções para os problemas cotidianos da fazenda, o que ficou conhecido como Pecuária de Precisão (PP) [Aquilani *et al.*, 2022]. A PP aborda a fazenda a partir de um ponto de vista analítico, permitindo a otimização de processos, redução de custos e melhorias na produção. Ainda, indicadores de produção podem ser usados para apoiar as decisões por meio de dados, medidos direta e continuamente dos animais [Fote *et al.*, 2020].

Os sistemas PP usam algoritmos de Aprendizado de Máquina (AM) para executar tarefas como previsão e classificação com base em dados obtidos de animais. Sensores de movimento, câmeras de vídeo, microfones, termômetros e similares são exemplos de dispositivos que podem ser usados para adquirir dados [Zhang *et al.*, 2021]. Diferentes tipos de algoritmos de AM são usados de acordo com os objetivos a serem alcançados, bem como considerando o tipo de dados produzidos pelos sensores [Sharma *et al.*, 2021] Random Forest (RF), Support Vector Machine (SVM), k-nearest neighbors (KNN), Classification and Regression Trees (CART), Linear Discriminant Análise (LDA) e variações, são exemplos dos algoritmos de AM mais usados para lidar com dados de sensores de movimento, e Deep Neural Networks (DNN) e variações são melhores para processamento de som e imagem [Oliveira *et al.*, 2021]

De acordo com [Neethirajan and Kemp, 2021], os principais desafios no monitoramento eficaz do bem-estar animal são: custo, a precisão e o tempo de discernimento. Apesar dos esforços crescentes na área, a maioria dos métodos disponíveis são demorados, trabalhosos e, portanto, caros. A precisão está relacionada à taxa de acerto

no diagnóstico proporcionado pelo sistema e o tempo de discernimento corresponde à latência entre o diagnóstico e o início da doença.

OBJETIVO:

No trabalho realizado por Silveira [da Silveira, 2021] foi desenvolvido um colar com sensores de movimento e temperatura para a coleta de dados do gado. Através do colar desenvolvido e da coleta de dados, o software constrói um etograma (registro de comportamentos exibidos do animal) em tempo real, tornando possível a identificação do seu comportamento como: *parado, andando, em pé, comendo, bebendo e deitado*. Por meio do sensor de temperatura e do etograma, o software pode classificar o animal entre saudável e doente.

O objetivo deste projeto consiste em criar etogramas utilizando-se de modelos de redes neurais recorrentes (Recurrent Neural Network, RNN) e de redes com memória de curto e longo prazo (Long Short-Term Memory, LSTM), empregando a mesma metodologia e dados abordados em [da Silveira, 2021].

METODOLOGIA

Para a realização do desenvolvimento do projeto, utilizou-se de dados coletados através de dois protótipos de colares colocados em duas novilhas saudáveis com aproximadamente 11 meses de idade. Os testes de campos foram executados na fazenda de pesquisa do Departamento de Zootecnia da Faculdade Luiz de Queiroz de Agronomia da Universidade de São Paulo (ESALQ/ USP) em outubro de 2021. Os dados foram registrados a cada segundo, e o termômetro realizava a medição a cada 20 minutos. Deste modo, coletou-se dados continuamente por cerca de 40 horas, de modo que, a cada movimentação das novilhas, que foram rotuladas e segmentadas em conformidade com seu comportamento naquele período. A rotulagem dos dados ocorreu de forma manual com o auxílio do vídeo do experimento.

No desenvolvimento do projeto utilizou-se majoritariamente de dois métodos de aprendizado de máquina, sendo eles: Redes Neurais Recorrentes (RNN) e Long-Short Term Memory (LSTM). A escolha dos métodos se deu principalmente por se tratarem de dados sequenciais, e esses serem intimamente relacionados a sequências e listas, assim a LSTM, é um método adequado para processar, prever e classificar séries temporais com intervalos de tempo indeterminado.

Os dados coletados pelos sensores fazem referência ao movimento em 3 eixos: X (movimentos horizontais), Y (movimentos longitudinais) e Z (movimentos verticais). Foram utilizados no cálculo de 8 variáveis, sendo as mesmas usadas como dados de entrada para métodos de aprendizado de máquina.

Os métodos de aprendizado de máquina RNN e LSTM foram comparados utilizando-se de uma amostragem de validação cruzada de 5 conjuntos. A capacidade preditiva de cada método foi avaliada considerando as seguintes métricas: precisão, sensibilidade, medida- F e o coeficiente Kappa (que compara a acurácia esperada com a observada).

RESULTADOS E DISCUSSÕES

No desenvolvimento das RNN's e LSTM's, utilizou-se 70% dos dados para treinamento, sendo divididos em 5 splits, a escolha se deu em decorrência dos testes realizados com esta porcentagem apresentarem valores melhores quando comparados aos testes com uma porcentagem menor de dados. Deste modo, optou-se por desenvolver quatro modelos de rede para a realização e compreensão dos testes.

Os modelos de rede a serem utilizados foram, R1: Uma rede neural recorrente com uma camada LSTM e a camada de saída; R2: Uma rede neural recorrente com uma camada LSTM, uma camada densa e a camada de saída; R3: Uma rede neural recorrente, uma camada LSTM bidirecional; R4: Uma rede neural recorrente com uma camada LSTM bidirecional, uma camada densa e uma camada de saída.

Para a realização dos testes para cada uma das redes desenvolvidas, utilizou-se dos modelos sendo passados determinados parâmetros em cada teste, assim, quando realizado o teste com os modelos passados, o teste com melhor resultado era exibido na saída. Os testes com melhores resultados foram explorados por meio da modificação de seus parâmetros, a modificação dos parâmetros do modelo se deu devido ao intuito de melhorar ainda mais os resultados dos testes.

Durante o desenvolvimento, pode-se perceber que, dentre os otimizadores utilizados nos testes, os que apresentaram melhor desempenho nas redes foram 'Adam' e 'RMSprop'. No que diz respeito a quantidade de neurônios utilizados, observou-se que uma menor quantidade (variando entre 25 à 100) apresentaram um melhor resultado, no que diz respeito às funções de ativação das redes, a função 'relu' apresentou melhores resultados quando comparado a 'sigmoid' e 'linear'.

A tabela abaixo contém alguns dos testes realizados para os modelos de rede desenvolvidos, selecionou-se os seis testes que apresentaram melhor desempenho para o treinamento do modelo. Criou-se um gráfico de dispersão para cada um dos modelos com os resultados dos testes, com o intuito de facilitar a visualização e entendimento. Vale-se ressaltar que, os testes realizados para o modelo de rede R4 não foram inseridos, pois eles ainda estavam sendo executados.

Desempenho dos testes no treinamento do modelo			
	R1	R2	R3
Teste 01	0,455416	0,443488	0,426103
Teste 02	0,454545	0,451048	0,44238
Teste 03	0,461948	0,452145	0,452732
Teste 04	0,464737	0,474018	0,442997
Teste 05	0,469120	0,472685	0,458262
Teste 06	0,429125	0,466249	0,450790

Tabela 1 – Resultado dos testes realizados para os modelos das redes desenvolvidas.

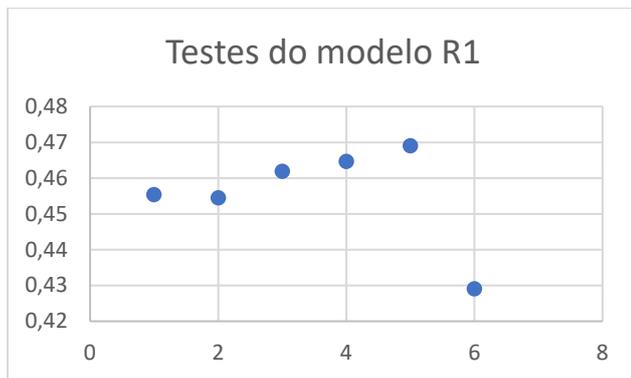


Gráfico 1: Gráfico de dispersão dos testes do modelo de rede R1

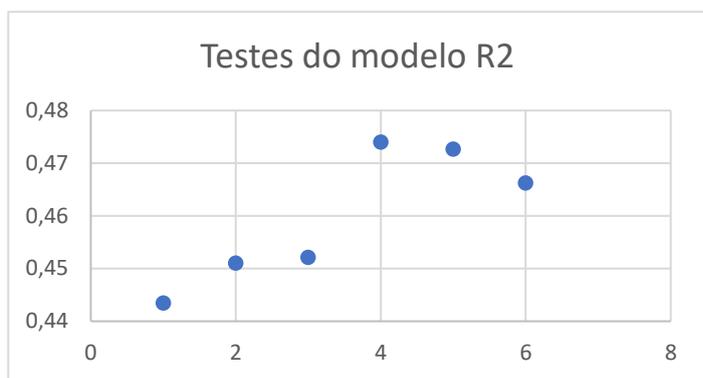


Gráfico 2: Gráfico de dispersão dos testes do modelo de rede R2

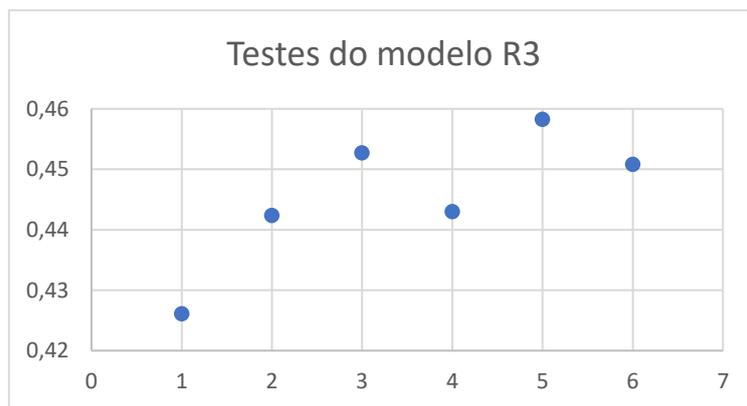


Gráfico 03: Gráfico de dispersão dos testes do modelo de rede R3

Até o atual momento, o melhor valor obtido no resultado dos testes foi de 0,474018, sendo este obtido no modelo R2, (uma rede neural recorrente com uma camada LSTM, uma camada densa e uma camada de saída). Espera-se, até o final do mês seguinte, obter resultados superiores a 0.5, já que desse modo, podemos obter um melhor desempenho das redes em relação aos conjuntos de dados utilizados no seu treinamento e conseqüentemente, uma maior confiabilidade no algoritmo.

No total, foram realizados cerca de 12 testes para cada modelo de rede. Os parâmetros e os modelos utilizados em cada teste, estão contidos no roteiro de testes, nesse, está descrito detalhadamente todos os testes realizados para cada rede. O roteiro de testes estará contido no anexo a ser entregue no relatório final, já que o mesmo documento apresenta cerca de 15 folhas no total.

CONCLUSÕES:

Com a realização do projeto, pode-se concluir que, as redes neurais recorrentes apresentam um bom desempenho em relação a dados de séries temporais. Os quatro modelos de redes desenvolvidos para experimentos apresentaram valores próximos. Por ser utilizada a mesma métrica em todos os modelos, sabe-se que a mesma pode ter limitado os valores obtidos no desempenho dos testes. Ainda assim, até o atual momento, o modelo de rede R2 apresentou o melhor desempenho, devendo-se ao fato de ser uma rede neural mais complexa, que tende a atuar com uma melhor precisão quando comparada aos outros dois modelos.

BIBLIOGRAFIA

[Aquilani et al., 2022] Aquilani, C., Confessore, A., Bozzi, R., Sirtori, F., and Pugliese, C. (2022). *Review: Precision livestock farming technologies in pasture-based livestock systems*. *Animal*, 16(1):100429

[Farooq et al., 2022] Farooq, M. S., Sohail, O. O., Abid, A., and Rasheed, S. (2022). *A survey on the role of iot in agriculture for the implementation of a smart livestock environment*. *IEEE Access*, 10:9483–9505.

[Fote et al., 2020] Fote, F. N., Roukh, A., Mahmoudi, S., Mahmoudi, S. A., and Debauche, O. (2020). *Toward a big data knowledge-base management system for precision livestock farming*. *Procedia Computer Science*, 177:136–142

[Neethirajan and Kemp, 2021] Neethirajan, S. and Kemp, B. (2021). *Digital livestock farming*. *Sensing and Bio-Sensing Research*, 32:100408.

[Oliveira et al., 2021] Oliveira, D. A. B., Pereira, L. G. R. Bresolin, T., Ferreira, R. E. P., and Dorea, J. R. R. (2021). *A review of deep learning algorithms for computer vision systems in livestock*. *Livestock Science*, 253.

[Ritter et al., 2017] Ritter, C., Jansen, J., Roche, S., Kelton, D. F., Adams, C. L., Orsel, K., Erskine, R. J., Benedictus, G., Lam, T. J., and Barkema, H. W. (2017). *Invited review: Determinants of farmers' adoption of management-based strategies for infectious disease prevention and control*. *Journal of Dairy Science*, 100(5):3329–3347.

[Sharma et al., 2021] Sharma, A., Jain, A., Gupta, P., and Chowdary, V. (2021). *Machine learning applications for precision agriculture: A comprehensive review*. *IEEE Access*, 9:4843–4873.

[da Silveira, 2021] da Silveira, L. R. (2021). *A livestock monitoring system for behavior and health condition classification*. Master's thesis, Faculdade de Tecnologia, Universidade Estadual de Campinas.

[Zhang et al., 2021] Zhang, M., Wang, X., Feng, H., Huang, Q., Xiao, X., and Zhang, X. (2021). *Wearable internet of things enabled precision livestock farming in smart farms: A review of technical solutions for precise perception, biocompatibility, and sustainability monitoring*. *Journal of Cleaner Production*, 312:127712