

PREDIÇÃO DO COMPORTAMENTO BOVINO COM SENSORES DE POSIÇÃO E MOVIMENTAÇÃO

Paulo Henrique Gonçalves Rezende^{1*}, Hana Karina Salles Rubinsztein²

1. Mestrando em Computação Aplicada da Universidade Federal de Mato Grosso do Sul (FACOM-UFMS)
2. Professora da FACOM-UFMS - Orientadora

Resumo

A Pecuária de Precisão é compreendida como gestão do rebanho com o uso da Tecnologia da Informação, definindo boas práticas na produção de carne. Dados de posicionamento, de automação e do comportamento do rebanho podem ser utilizados em sistemas que permitem reduzir os custos da produção e os impactos ambientais, e aumentar a produtividade.

Este trabalho propõe a identificação do comportamento bovino, utilizando colares com sensores de movimentação e posicionamento aplicados nos animais, para registrar sua movimentação. Em seguida, algoritmos de Aprendizagem de Máquina supervisionada são usados para definir o modelo que fará a predição do comportamento. Essa pesquisa conta com o apoio e parceria entre Embrapa Gado e Corte e FACOM-UFMS.

Palavras-chave: pecuária de precisão; aprendizado de máquina; pré-processamento.

Introdução

A agropecuária é a principal atividade econômica do Brasil. O MAPA (Ministério da Agricultura, Pecuária e Abastecimento) indicou que a Agropecuária em 2017 obteve um crescimento no PIB de 13,4% em relação a 2016 [1]. A Pecuária de Precisão é um caminho para atingir as metas das projeções, pois pratica a gestão do gado com TI (Tecnologia da Informação), gerando boas práticas na produção de couro e carne [2], ajudando o pecuarista com sistemas de monitoramento e comportamento animal, sem agredir o meio ambiente e gerar custos excessivos.

A Embrapa Gado de Corte e a FACOM-UFMS possuem projetos de desenvolvimento de tecnologias para a Pecuária de Precisão. Como o gerenciamento tradicional, onde um indivíduo observa o gado e anota informações, é considerada uma atividade limitada e onerosa, projetos de gestão de precisão usam recursos com mais eficiência, utilizando sensores de monitoramento para registrar comportamentos bovinos e enviá-los para análise do pecuarista [3].

Este trabalho emprega sensores de movimentação e posicionamento para inferir dados de comportamento animal, aumentando o nível da gestão de precisão. Para essa inferência, são utilizadas ferramentas de aprendizagem de máquina que determinarão o comportamento bovino, aprimorando o tratamento dos dados que poderão influenciar as escolhas de manejo e as tomadas de decisão.

Metodologia

Nesse trabalho foram desenvolvidos colares com sensores de movimentação e de posicionamento que são dispostos individualmente aos animais. Esses sensores são tecnologias que simplificam o gerenciamento do gado e sistemas que utilizam desses recursos conseguem destacar o potencial de análise do monitoramento bovino [3]. No experimento a campo para coleta dos dados, os animais foram expostos ao pasto, onde agem naturalmente enquanto os sensores registram as informações e um observador registra o comportamento através de um *software* de anotação. As anotações são as classes utilizadas como rótulos na aprendizagem de máquina nas etapas subsequentes. As classes observadas foram: Comendo, Ruminando (Deitado), Ruminando (Em pé), Deitado, Andando e Em pé.

Em seguida os dados foram dispostos para a execução do pré-processamento. Como alguns dados se encontram com informações defeituosas, esse trabalho propôs a criação de 3 cenários para o aprendizado de máquina, sendo eles o **Cenário Original**, onde apenas se desconsidera os dados defeituosos; o **Cenário Expandido**, que substitue esses registros por uma média aritmética dos dados equidistantes; e o **Cenário Reduzido**, proposto para usar apenas dados próximos das transições de comportamento executadas pelos bovinos. Para multiplicar a quantidade de cenários, foi empregada a metodologia dos trabalhos de Jesus [4], que usou apenas registros de posicionamento com informações que derivaram deste, e Lomba [5], que uniu o posicionamento com os dados de movimentação, bem como a união das duas, a fim de melhorar o treinamento dos modelos de aprendizagem.

Posteriormente é utilizada a ferramenta de aprendizagem de máquina, o *Scikit-Learn* [6], para desenvolver o modelo de aprendizagem. Os dados pré-processados são embaralhados de maneira aleatória, mantendo sua origem temporal, para simular a maneira que uma futura aplicação irá trabalhar, concebendo ao algoritmo de aprendizagem um cenário próximo ao real, melhorando a taxa de aprendizagem do modelo.

Além de manter os dados originais, existem 2 métodos que refinam os dados que podem gerar melhores resultados, sendo a **Normalização**, na qual opera uma escala de amostras dos dados de forma individual, tendo uma norma unitária como resultado; e a **Padronização**, em que realiza cálculos de desvio padrão e média com os dados. Ambos os métodos possuem resultados mais simples do que os dados brutos originais, tornando-as outras metodologias para o encontro do melhor modelo de aprendizagem desses dados. A forma utilizada para o aprendizado é a supervisionada, pois já existe a saída, que são as classes definidas pelo *software* de anotação,

bastando o resultado estar dentro desse conjunto para caracterizar um acerto ou erro.

Os modelos com os melhores resultados serão empregados nas futuras coletas, que gerarão uma aplicação que fará a classificação supervisionada para as novas bases de dados.

Resultados e Discussão

Foram utilizadas 3 bases de dados para o aprendizado de máquina. O modelo de aprendizagem foi embaralhado utilizando os métodos *Cross-Validation* e o *GridSearch* [7], esse último também escolhe os melhores parâmetros de acordo com o *dataset* usado. Assim, os algoritmos de classificação supervisionada empregados foram o *Logistic Regression*, um modelo linear de classificação, e o *Decision Tree*, um conceito de árvore de decisão. A Tabela 1 ilustra os resultados das acurácias, ou seja, a porcentagem total de acertos para cada algoritmo e cenário proposto.

Base Jesus: Cenário Original				
Algoritmo	Dado Original	Dado Padronizado	Dado Normalizado	Média
Logistic Regression	45%	43%	43%	44%
DecisionTree Classifier	37%	40%	41%	39%
Média	41%	42%	42%	42%
Base Lomba: Cenário Original				
Algoritmo	Dado Original	Dado Padronizado	Dado Normalizado	Média
Logistic Regression	43%	41%	42%	42%
DecisionTree Classifier	36%	42%	46%	41%
Média	40%	42%	44%	42%
Base Jesus+Lomba: Cenário Original				
Algoritmo	Dado Original	Dado Padronizado	Dado Normalizado	Média
Logistic Regression	43%	41%	42%	42%
DecisionTree Classifier	34%	43%	39%	39%
Média	39%	42%	41%	41%

Tabela 1: Resultados das acurácias do Cenário Original

Entende-se que a média aritmética ficou em 42%, sendo que a junção das metodologias de Jesus [4] e Lomba [5] foi o pior cenário, constando que a sua união apresenta mais ruídos do que melhorias para o Cenário Original. Isoladamente, é notório observar que os dados padronizados apresentaram a melhor acurácia dentre as 3 bases, com a única exceção da Base Lomba [5] com os dados Normalizados, que apresentou a média de 44% nos dois algoritmos. Na comparação com os algoritmos, o *Logistic Regression* alcançou a média de 44% no cenário Jesus [4]. Se estabelece que as oscilações de cada algoritmo, com cada cenário, bem como as mudanças de dados, não determinaram um destaque no Cenário Original.

O próximo passo era executar no Cenário Reduzido. A Tabela 2 informa os resultados que foram obtidos. Nota-se que, em comparação com o cenário anterior, esse apresentou um agravamento de 22%, indicando que o intuito de reduzir os dados apenas nas transições de comportamento foi ineficaz.

Base Jesus: Cenário Reduzido				
Algoritmo	Dado Original	Dado Padronizado	Dado Normalizado	Média
Logistic Regression	25%	11%	11%	16%
DecisionTree Classifier	22%	22%	27%	24%
Média	24%	17%	19%	20%
Base Lomba: Cenário Reduzido				
Algoritmo	Dado Original	Dado Padronizado	Dado Normalizado	Média
Logistic Regression	26%	11%	11%	16%
DecisionTree Classifier	18%	22%	21%	20%
Média	22%	17%	16%	18%
Base Jesus+Lomba: Cenário Reduzido				
Algoritmo	Dado Original	Dado Padronizado	Dado Normalizado	Média
Logistic Regression	26%	11%	11%	16%
DecisionTree Classifier	23%	22%	18%	21%
Média	25%	17%	15%	19%

Tabela 2: Resultados das acurácias do Cenário Reduzido

Por fim, foi executado os algoritmos no Cenário Expandido. Nota-se que foi o Cenário que apresentou a melhor acurácia média dentre os outros, demonstrando que incluir as médias de dados corrompidos foi promissora, o que fez a ferramenta de aprendizado de máquina aprender com mais facilidade. De forma isolada,

o Cenário de Lomba [5] foi o pior resultado com 47% dos acertos, enquanto Jesus [4] demonstrou supremacia com seus 49%. Dentre as mudanças de dados, podemos verificar que os dados originais apresentaram a melhor taxa de acertos na Base Lomba e Jesus+Lomba, enquanto a Base Jesus esse ficou empatado com os dados padronizados. Por fim, o *Decision Tree* foi o melhor algoritmo em relação ao *Logistic Regression*. Esses resultados podem ser conferidos na Tabela 3.

Base Jesus: Cenário Expandido				
Algoritmo	Dado Original	Dado Padronizado	Dado Normalizado	Média
Logistic Regression	47%	48%	47%	47%
DecisionTree Classifier	54%	53%	47%	51%
Média	51%	51%	47%	49%
Base Lomba: Cenário Expandido				
Algoritmo	Dado Original	Dado Padronizado	Dado Normalizado	Média
Logistic Regression	47%	47%	47%	47%
DecisionTree Classifier	52%	45%	43%	47%
Média	50%	46%	45%	47%
Base Jesus+Lomba: Cenário Expandido				
Algoritmo	Dado Original	Dado Padronizado	Dado Normalizado	Média
Logistic Regression	47%	47%	47%	47%
DecisionTree Classifier	50%	49%	45%	48%
Média	49%	48%	46%	48%

Tabela 3: Resultados das acurácias do Cenário Expandido

Conclusões

Essa pesquisa visou trabalhar analisar o comportamento bovino utilizando ferramentas como Aprendizado de Máquina para poder inferir com mais exatidão os resultados. Pode-se observar que, ao embaralhar os dados, simulando um experimento real no campo, mesmo apresentando melhorias, ainda está longe de possuir uma base de informação consisa e adequada para uma aplicação real da Embrapa Gado de Corte. Os trabalhos futuros executarão novas experiências com as bases, incluir novos dados e executar o mesmo procedimento com outros algoritmos com a expectativa de que os resultados sejam aperfeiçoados.

Esse trabalho celebrou mais uma parceria entre a Embrapa Gado de Corte e a FACOM-UFMS, buscando novas soluções para a Pecuária de Precisão, em especial na classificação dos comportamentos bovinos e seus padrões de comportamento.

Referências bibliográficas

- [1] MAPA. <Projeções do Agronegócio>. 2017. Endereço:<http://www.agricultura.gov.br/assuntos/politica-agricola/todas-publicacoes-de-politica-agricola/projecoes-do-agronegocio/projecoes-do-agronegocio-2017-a-2027-versao-preliminar-25-07-17.pdf>.
- [2] CÁCERES, E. N. et al. <Computational precision livestock – position paper>. In: II Workshop of the Brazilian Institute for Web Science Research. Rio de Janeiro, 2011. Endereço:http://www.gpec.ucdb.br/pistori/publicacoes/caceres_webscience2011.pdf.
- [3] SMITH, D. et al. <Behavior classification of cows fitted with motion collars: Decomposing multi-class classification into a set of binary problems>. Computers and Electronics in Agriculture, v. 131, p. 40–50, dez. 2016. Endereço:<http://www.sciencedirect.com/science/article/pii/S0168169916303180ab005>.
- [4] JESUS, L. de. <Identificação do Comportamento Bovino por meio do Monitoramento Animal>. Dissertação — UFMS, jul. 2014. Endereço:<http://repositorio.cbc.ufms.br:8080/jspui/handle/123456789/2075>.
- [5] LOMBA, L. F. D. <Identificação do Comportamento Bovino a partir dos Dados de Movimentação e do Posicionamento do Animal>. Dissertação — UFMS, dez. 2015. Endereço:<http://repositorio.cbc.ufms.br:8080/jspui/handle/123456789/2627>.
- [6] PEDREGOSA, F. et al. <Scikit-learn: Machine learning in python>. Journal of Machine Learning Research, v. 12, p. 2825–2830, 2011. Endereço:<http://www.jmlr.org/papers/v12/pedregosa11a.html>.
- [7] BERGSTRA, J.; BENGIO, Y. <Random search for hyper-parameter optimization>. Journal of Machine Learning Research, v. 13, 2012. Endereço:<http://www.jmlr.org/papers/volume13/bergstra12a/bergstra12a.pdf>.