

Classificação do comportamento do gado leiteiro usando dados de sensores de movimentos com redes Multilayer Perceptron

Anderson da Silva Santos
Departamento de Estatística e Informática, DEINFO
Universidade Federal Rural de Pernambuco
Recife, Pernambuco
anderson.silvasantos@ufrpe.br

Resumo—Este trabalho consiste no desenvolvimento de uma rede Multilayer Perceptron (MLP) para classificação do comportamento do gado leiteiro por meio de dados de sensores (acelerômetro, giroscópio e magnetômetro) embutidos em uma coleira. Os comportamentos analisados nesse estudo são pastando, comendo, deitado e em pé. Foram avaliadas algumas topologias de rede para avaliação da performance na acurácia. Uma comparação dos resultados, com e sem normalização dos dados, foi verificada. O melhor valor da acurácia foi obtido na configuração com a maior quantidade de neurônios na camada intermediária.

Index Terms—redes neurais, aprendizagem de máquina, multilayer perceptron, pecuária de precisão, etologia, aprendizagem supervisionada, séries temporais.

I. INTRODUÇÃO

O crescimento do acesso a tecnologia é um dos principais fatores por trás da pecuária de precisão, onde o animal não é mais monitorado em nível de rebanho, mas sim, em um nível individual. Como é impraticável para os agricultores observarem continuamente cada animal do rebanho, foram desenvolvidos sistemas para monitorar automaticamente o comportamento de cada animal.

Esses tipos de sistemas abrem espaço para novas oportunidades para o monitoramento dos comportamentos, ações, produtividade e bem-estar de cada animal. Uma aplicação importante baseada nesses comportamentos inferidos é a previsão do momento ideal para a inseminação em uma vaca, identificando onde no ciclo estral há um aumento significativo na intensidade da atividade.

Para discriminar entre classes de comportamento, o sistema precisa capturar padrões exclusivos para cada classe. A escolha dos sensores e o local de colocação desses sensores nos animais limitam os tipos de movimento que podem ser representados.

Além da captura, o pré-processamento desses dados se torna uma etapa necessária para a remoção de ruídos presentes durante a captura nos sensores. Há a possibilidade dessa etapa já ocorrer no próprio dispositivo com a aplicação de filtros. Modelos estatísticos são usados para o processamento dos dados originados por sensores em comportamentos significativos.

Um detalhe sobre os comportamentos desses animais é que esses movimentos geralmente são contínuos e repetitivos, ou seja, não há mudanças drásticas em um espaço de tempo curto no comportamento do animal.

II. BASE DOS DADOS

A base de dados foi gerada por um trabalho de dissertação de mestrado desenvolvido na Universidade Federal de Mato Grosso do Sul (UFMS) com parceria da Embrapa Gado de Corte, na cidade de Campo Grande, Mato Grosso do Sul.

A seguir alguns dados sobre os experimentos:

- 2 períodos: o primeiro de 25/03/15 a 30/03/15. O segundo de 06/04/15 a 09/04/15
- Animais: 4 vacas
- Dispositivo: coleira
- Foram utilizadas 4 coleiras (A, B, C e D). Entretanto, houve algumas falhas durante o experimento e, no final, ficaram as seguintes bases:
 - A2 e A3
 - B2 e B3
 - C3 e C4
 - D1, D2, D3 e D4
- Sensores: acelerômetro, giroscópio e magnetômetro
- As coletas foram feitas na frequência de 1 Hz, ou seja, a cada 1 segundo eram capturados os dados de todos os sensores
- Os dados foram armazenados em um cartão SD para posteriormente serem carregados e processados
- Classes dos comportamentos: pastar, em pé, deitado e andando. A seguir uma descrição de cada um dos comportamentos
 - Pastando: caracterizado pelo animal sobre as quatro patas, com a cabeça baixa procurando ou mastigando o capim. O animal pode ou não estar em movimento, já que ele pode estar se deslocando à procura de capim;
 - Andando: o animal também está sobre as quatro patas, porém com o pescoço reto (apontando o focinho para frente) e se deslocando pela área de pasto;
 - Em Pé: o animal está sobre as quatro patas, com a cabeça erguida e não há deslocamento;
 - Deitado: o animal está com as patas abaixadas e com a barriga tocando o solo.

Por fim, foi gerado um arquivo *csv* o qual contém todos os dados aglutinados que foram utilizados.

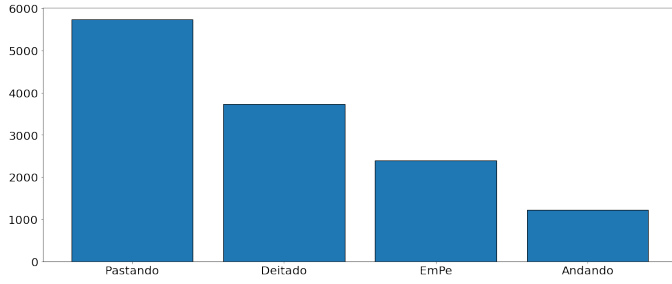


Figura 1: Distribuição dos dados por classe.

III. VISUALIZAÇÃO DOS DADOS

Antes de partir para a implementação da MLP, foi realizado uma série de scripts para a visualização dos dados. Esta etapa é importante para o trabalho, pois permite que os dados, junto com a sua estrutura, possam ser melhor compreendidos e consequentemente ajudar na tomada de decisão de quais técnicas utilizar para o processamento.

A Figura 1 apresenta a distribuição dos dados por cada uma das classes. É possível perceber uma forte desproporcionalidade entre as amostras da classe pastando e andando. Isso é um ponto importante a se destacar, pois pode influenciar significativamente no treinamento da rede e consequentemente apresentar um sobreajuste nos dados e a perda da capacidade de generalização da rede.

A Figura 2 mostra os histogramas de cada coleira para cada sensor. Essa outra forma de visualização nos mostra que os dados de diferentes coleiras, apesar de possuir os mesmos sensores, apresentam diferenças entre si. Essa complexidade é comum nesse contexto, pois variáveis como o ruído na captura dos dados em cada sensor, a posição da coleira em cada animal e a raça causam alterações na captura dos dados. Há trabalhos também que relatam mudanças no comportamento de acordo com a estação do ano.

A Figura 3 ilustra as séries temporais de cada sensor em relação à respectiva coleira. É notório uma grande diferença na representabilidade entre as coleiras. Por exemplo, a coleira A no turno 3 contém mais dados do que a coleira D no turno 4. Tal superioridade numérica também afetará o desempenho do classificador por meio de um viés para uma respectiva coleira.

Portanto, essas variáveis somadas resultam em uma maior complexidade para uma solução robusta nesse contexto da pecuária de precisão. A Figura 4 aglutina todos as séries temporais de todas as coleiras.

IV. ARQUITETURA

A. Parâmetros da rede

- Função de ativação: regressão logística (sigmóide)
- Algoritmo de otimização: gradiente descendente
- Termo de regularização L2: 0.0001
- Atualização dos pesos em lotes com tamanho de 24 amostras
- Taxa de aprendizagem: 0.3
- Número de épocas: 200

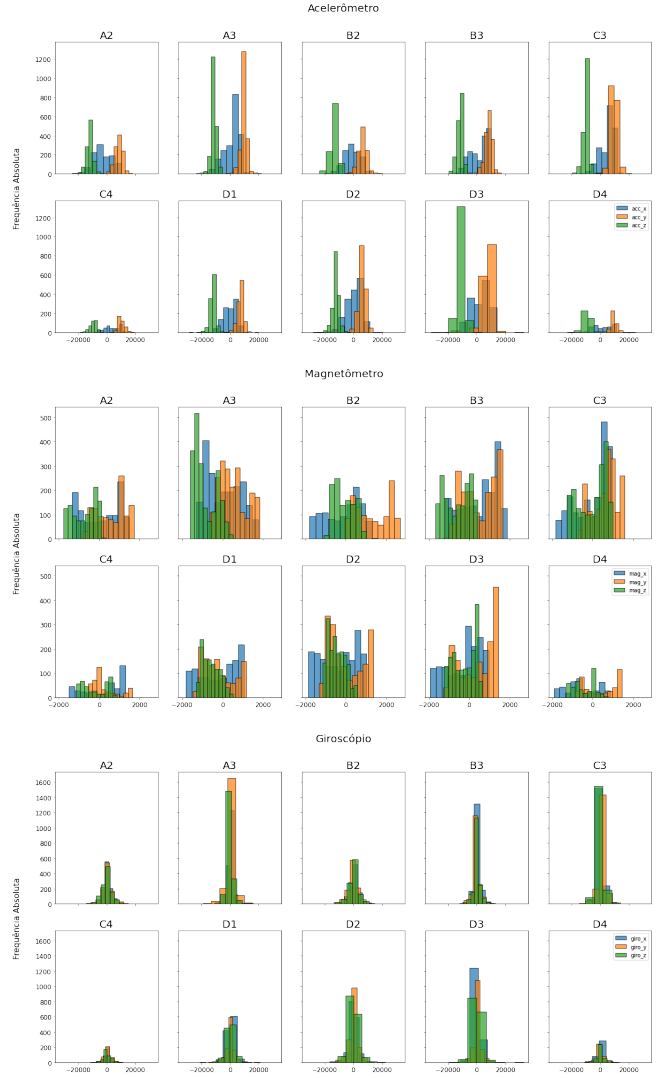


Figura 2: Histogramas por coleira.

- Randomização das amostras para cada iteração
- Momentum: 0.9

B. Camada Oculta

Foram realizadas algumas combinações em relação à quantidade de neurônios na camada oculta. Os seguintes valores foram definidos para a topologia da camada oculta:

- Primeiro nível: 9, 18, 27 e 36 neurônios
- Segundo nível: 6, 8, 10 e 12 neurônios

V. VALIDAÇÃO

Para validação, foi utilizada a técnica de validação cruzada com o valor do k igual a 10. A Figura 6 ilustra a divisão dos dados para essa abordagem. O valor da acurácia média foi utilizado como medida de avaliação para comparações entre os conjuntos de valores dos parâmetros da rede.

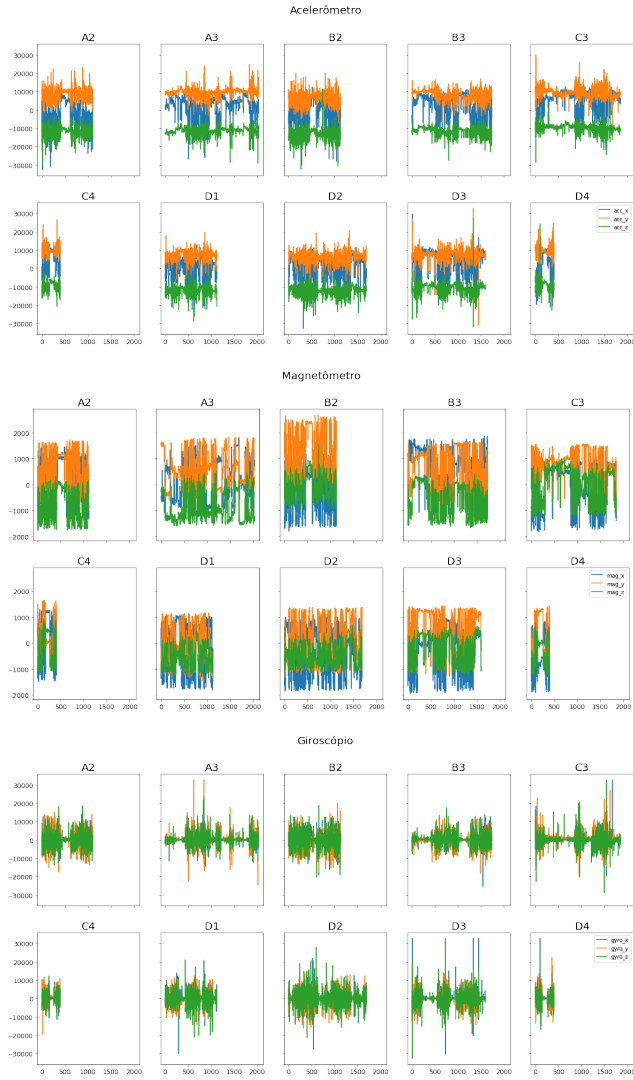


Figura 3: Distribuição dos dados por coleira.

VI. RESULTADOS

Os resultados dos experimentos são mostrados nas Tabelas I e II. Além da avaliação da acurácia em relação à variação da topologia da rede na camada oculta, foi realizada uma comparação com e sem o processo de normalização aplicado aos dados. Esse processo de normalização foi feito de acordo com a Equação 1.

$$z = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (1)$$

Onde z é a variável normalizada; x é a variável original e as funções \min e \max retornam o menor e o maior valor de cada característica presente na base respectivamente. Essa técnica foi utilizada para fazer com que os dados estivessem na mesma escala e não houvesse nenhum ganho de vies apenas pela magnitude dos valores das características. Além do mais, esse procedimento permite que o gradiente descendente possa convergir mais rápido.

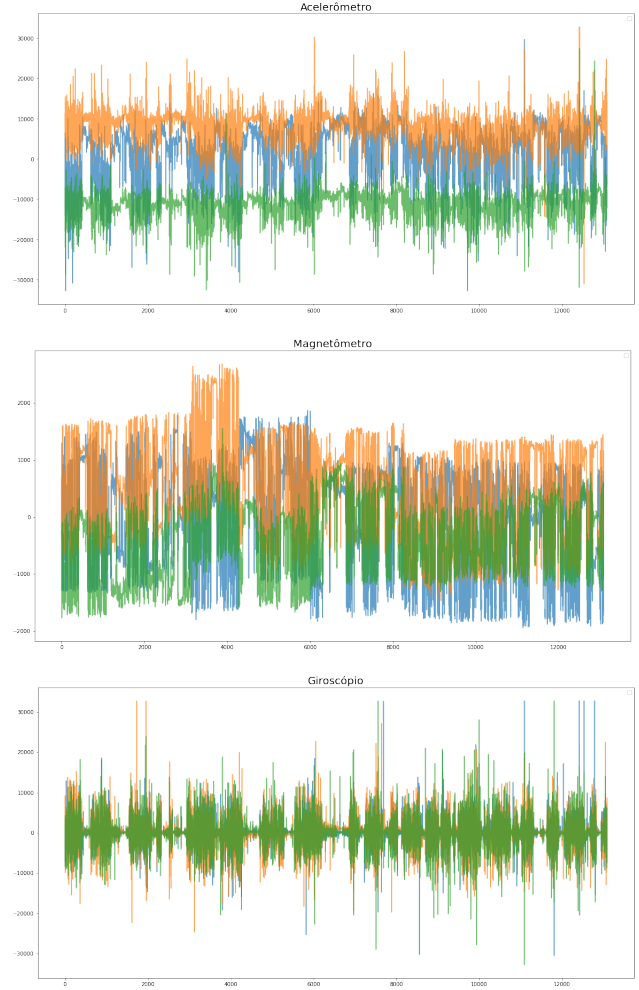


Figura 4: Séries temporais aglutinadas.

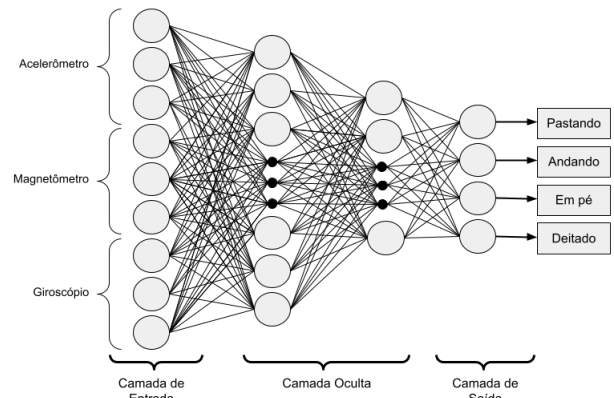


Figura 5: Arquitetura da rede MLP.

Base de Dados										
Treinamento							Validação			
n=1	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
n=2	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
n=3	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
n=4	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
n=5	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
n=6	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
n=7	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
n=8	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
n=9	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
n=10	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10

Acurácia Média

Figura 6: Validação cruzada com o número de grupos igual a 10.

A tabela I mostra os resultados sem o processo de normalização. É possível perceber uma baixa acurácia no processo de validação. Uma boa hipótese para tais valores, é a não convergência no processo de treinamento do gradiente descendente com as 200 épocas.

1º nível	2º nível	Acurácia (%)
9	6	41.66 (0.0674)
18	8	40.89 (0.0676)
27	10	44.22 (0.0149)
36	12	44.55 (0.0642)

Tabela I: Resultados das acurácias das redes sem a normalização dos dados. Os valores entre parênteses são os desvios-padrão entre os 10 folds.

Já a Tabela II mostra os resultados com o processo de normalização. Aqui temos uma melhora significativa na acurácia. Nesse caso, o gradiente descendente, com o mesmo número de épocas, conseguiu convergir para uma área mais próxima do valor do mínimo global.

1º nível	2º nível	Acurácia (%)
9	6	76.32 (0.0707)
18	8	76.25 (0.0728)
27	10	77.82 (0.0531)
36	12	78.03 (0.0597)

Tabela II: Resultados das acurácias das redes com a normalização dos dados. Os valores entre parênteses são os desvios-padrão entre os 10 folds.

VII. CONCLUSÃO

Neste trabalho, foi realizado um estudo com diferentes topologias na camada oculta de uma rede MLP para a classificação do comportamento do gado leiteiro usando dados de sensores de movimento embutidos em uma coleira.

A construção de uma série de gráficos possibilitou as visualizações e posteriormente uma maior compreensão das estruturas dos dados. Ficaram evidentes os desafios da captura e rotulação dos dados que são aderentes a esse contexto da pecuária de precisão.

Uma validação cruzada foi utilizada para avaliação da acurácia dos modelos. Comparações entre as acurácias, com e sem normalização, foram realizadas. Pode-se confirmar uma maior convergência do gradiente descendente após a normalização dos dados e consequentemente uma melhoria significativa na acurácia. Portanto, a normalização é uma das técnicas fundamentais para melhor performance nesse contexto desse tipo de rede.

Além do mais, a melhor configuração da rede foi no caso em que houve a maior quantidade de neurônios na camada intermediária com 36 e 12 neurônios no primeiro e segundo nível respectivamente. Esse aumento na quantidade de neurônios permitiu que a rede possuísse maior robustez na generalização no processo de classificação.

VIII. TRABALHOS FUTUROS

Algumas idéias precisam ser avaliadas na continuidade desse trabalho. A seguir uma lista de propostas que poderão ser analisadas em trabalhos futuros.

- Aumentar a quantidade de neurônios na camada intermediária e verificar a resposta da rede em relação à acurácia e a presença de overfitting.
- Aplicar um pré-processamento nos dados (filtros passa alta/baixa)
- Aumentar a quantidade de épocas no treinamento
- Utilizar uma abordagem evolucionária para obter os melhores parâmetros da rede
- Testar outros algoritmos além do gradiente descendente como o LBFGS e ADAM
- Utilizar outras funções de ativação como por exemplo a tangente hiperbólica, Relu, Elu e entre outras