



## MODELAGEM E IMPLEMENTAÇÃO DO SISTEMA ESTAÇÃO CAVALO PANTANEIRO PARA ANÁLISE DAS CARACTERÍSTICAS MORFOMÉTRICAS

*B.R. de Barros<sup>1</sup>, E.L. dos Santos<sup>1</sup>, S.A. Santos<sup>1</sup>, F.F.de Souza<sup>1</sup>*

(1) Instituto Federal do Mato Grosso do Sul, Rua Delamare, 79304-000, Corumbá, MS, brunnoruiz@hotmail.com, eduardolourenco\_1@hotmail.com, fabricia.souza@ifms.edu.br

(2) Embrapa Pantanal, Rua 21 de Setembro, 1880, 79320-900 Corumbá, MS, sandra.santos@embrapa.br

**Resumo:** A caracterização morfofuncional do cavalo Pantaneiro é um dos critérios de conservação, seleção e melhoramento da raça. Visando extrair conhecimentos e padrões de 15 medidas lineares comumente tomadas no registro da raça, este estudo objetivou implementar sistema de dados denominado “Estação Cavallo Pantaneiro” (SDECP). O modelo conceitual criado para a construção do SDECP baseou-se na análise, modelagem, implementação, teste e aplicação. Para a modelagem do banco de dados, utilizou-se a ferramenta brModelo, para mineração a ferramenta WEKA, para implementação a (IDE) NetBeans e o MATLAB. Para a extração de informações e formulação de padrões por meio de dados morfométricos da raça e índices gerados, utilizou-se da árvore de decisão e lógica *fuzzy*. Com o SDECP criado foi possível explorar os dados morfométricos por meio de gráficos reunidos ou individualizados, assim como definir padrões da raça. Este banco mostrou ampla gama de aplicação nas tomadas de decisão referente às medidas morfofuncionais da raça.

**Palavras-chave:** banco de dados, inteligência artificial, lógica fuzzy, mineração de dados, raças localmente adaptadas.

### MODELLING AND IMPLEMENTING OF SYSTEM CAVALO PANTANEIRO STATION FOR THE ANALYSIS OF MORPHOMETRIC MEASURES

**Abstract:** The morphofunctional characterization of Pantaneiro horses is one of the criteria for conservation, selection and genetic improvement of the breed. Aiming to extract knowledge and patterns of 15 measures commonly taken during the registration of the breed, a database named “Cavallo Pantaneiro station” (DBCPS) was implemented. The conceptual model to the construction of the DBCPS proceeded from analysis, modelling, implementation, test and application. The following tools were adopted: BR Model for modelling the database, WEKA for data mining, (IDE) NetBeans for implementation and the MATLAB. In this study was adopted decision tree and fuzzy logic to extract information and patterns formulations from morphometric data and indices. With the SCPD was possible to explore the morphometric data through graphs pooled or individualized as well to define patterns of the breed. This system showed wide application range in decision-making concerning the use of morphofunctional measures of the breed.

**Keywords:** data base, artificial intelligence, fuzzy logic, data mining, locally adapted breeds.

## 1. Introdução

O cavalo pantaneiro é uma raça localmente adaptada ao Pantanal, fundamental no manejo extensivo do gado de corte efetuado nas fazendas do Pantanal. Além da funcionalidade na lida do gado, o cavalo também tem se destacado nas provas equestres como a do Laço Comprido, características estas que tem valorizado a raça comercialmente (SANTOS et al., 2003). Esta valorização da raça tem aumentado o interesse na sua criação, consequentemente tem aumentado o número de registros da raça que conta atualmente com aproximadamente 5000 cavalos com registros definitivo. No registro dos animais, a ABCCP (Associação Brasileira de Criadores de Cavalos Pantaneiros) utiliza medidas morfométricas. Estas mesmas medidas também vem sendo monitoradas no núcleo de conservação da Embrapa Pantanal desde 1988, gerando uma grande quantidade de dados. De maneira geral, essa massa de dados é pouco explorada nas tomadas de decisão devido à dificuldade de manipular os dados.

Para auxiliar na construção de um banco de dados flexível às necessidades locais e imediatas dos usuários, há a necessidade da construção de um BD (Banco de dados) local associado com técnicas de inteligência artificial que estuda como representar ou formalizar esses fatos em um formato que seja “tratável” pelo computador, denominado de Representação do Conhecimento (RC) (ALVARENGA, 2003). No BD essa abstração busca que os usuários consigam adequar os dados às informações e conhecimentos gerados pelos mesmos (WOLPP, 2004). O banco de dados obedece a uma arquitetura entre cliente e servidor, distribuída assim nesses dois módulos. (ELMASRI e NAVATHE, 2011). Assim, o processo futuro decisório referencial, trata da mineração desses dados,

sendo fundamental, então, análise de modelos e índices previamente estabelecidos, como o uso conceitual da Inteligência Artificial.

Este estudo objetivou implementar um Sistema com apoio gerencial SQL (HeidiSQL – Software gratuito) denominado sistema de dados “Estação Cavalos Pantaneiros” (SDECP) para auxiliar no processo de extração de informações (padrões) de dados morfométricos da raça através da árvore de tomada de decisão binária e lógica *fuzzy*, baseando-se na mineração dos dados de entrada e na padronização de parâmetros de decisão sistêmica.

## 2. Materiais e Métodos

A modelagem do sistema de dados Estação Cavalos Pantaneiros (SDECP) seguiu três etapas: modelagem conceitual, projeto lógico e projeto físico.

### 2.1. Desenvolvimento do modelo conceitual

O modelo conceitual seguiu o organograma de execução: análise, modelagem e implementação, teste e aplicação. Primeiramente definiu-se com os usuários a necessidade de um sistema para gerenciar as informações do Banco de dados morfométricos lineares utilizados para caracterizar fenotipicamente os cavalos do Núcleo de criação de cavalos Pantaneiros da Embrapa Pantanal e também nos registros de cavalos efetuados pela ABCCP (Associação Brasileira de Criadores de Cavalos Pantaneiros). Os dados morfométricos referem-se a 15 medidas lineares: altura da cernelha (AC), altura do dorso (AD), altura da garupa (AG), altura dos costados (ACost), comprimento da cabeça (CC), comprimento do pescoço (CP), comprimento dorso-lombar (CD), comprimento da garupa (CG), comprimento da espádua (CE), comprimento do corpo (Ccorp), largura da cabeça (LCD), largura do peito (LP), largura da anca (LA), perímetro torácico (PT) e perímetro da canela (PC). As medidas lineares utilizadas neste estudo constam na Tabela 1.

Tabela 1. Medidas morfométricas do cavalo Pantaneiro (SANTOS, 2003)

VARIÁVEIS OBSERVADAS	MACHOS			FÊMEAS		
	N	Média ± s	CV(%)	N	Média ± s	CV
<b>Alturas (cm):</b>						
Cernelha (AC)	180	141,7 ± 1,7	0,01	1035	136,9 ± 2,2	0,02
Dorso (AD)	180	138,2 ± 2,6	0,02	1032	134,5 ± 2,7	0,02
Garupa (AG)	180	141,8 ± 2,2	0,01	1034	137,9 ± 2,5	0,02
Costado (AT)	180	54,4 ± 2,5	0,05	1031	58,5 ± 3,5	0,06
<b>Comprimentos (cm):</b>						
Cabeça (CC)	180	54,4 ± 2,5	0,05	1034	53,6 ± 2,5	0,05
Pescoço (CP)	180	57,7 ± 4,9	0,08	1034	57,5 ± 6,0	0,10
Dorso (CD)	147	50,4 ± 8,1	0,16	914	42,6 ± 9,6	0,23
Garupa (CG)	180	45,2 ± 2,0	0,06	1034	44,9 ± 2,8	0,06
Espádua (CE)	180	47,6 ± 3,6	0,08	1034	46,2 ± 3,4	0,07
Corpo (CR)	180	146,2 ± 22,4	0,15	1033	141,3 ± 5,1	0,04
<b>Larguras (cm):</b>						
Cabeça (LC)	180	20,9 ± 2,0	0,10	1033	20,5 ± 2,5	0,12
Peito (LP)	180	34,1 ± 3,1	0,09	1034	31,6 ± 3,0	0,09
Anca (LA)	180	43,5 ± 2,7	0,06	1034	43,5 ± 3,0	0,07

Destas medidas lineares foram calculados índices morfométricos como o índice corporal que refere-se entre a relação de Ccorp e PT (Tabela 2).

Tabela 2. Exemplo de classificação de índice corporal para definição de aptidão funcional com auxílio do *fuzzy*.

Índice Corporal (IC=Ccorp/PT) (1)	
IC ≥ 90	Animal longilíneo com maior aptidão para velocidade (corrida)
85 ≤ IC ≤ 90	Animal mediolíneo, apto para sela (padrão esperado da raça).
IC ≤ 85	Animal brevilíneo, apto para tração.

$Ccorp$  = Comprimento do Corpo,  $PT$  = Perímetro do Tórax.

Na análise de software foram escolhidas as ferramentas mais adequadas a serem utilizadas no processo de análise do programa. Para a modelagem do BD, utilizou-se a ferramenta brModelo, para gerenciamento do BD o HeidiSQL, para mineração a ferramenta WEKA, para implementação a IDE Netbeans (utilizando para programação a linguagem Java) e o MATLAB, ferramenta matemática para funcionalidades de cálculos variados. O processo de análise de dados visou a identificação dos atributos (variáveis) para manipulação do sistema, os atributos que compõem as medidas morfométricas, os atributos relacionados com o cavalo (registro, nome, data de nascimento, origem, proprietário), com o criatório, assim como os cálculos para estimativa do peso corporal e de índices de avaliação morfométricas (SANTOS et al., 2003; McMANUS et al., 2008). No desenvolvimento do banco de dados, construiu-se um modelo inicial por meio da ferramenta BrModelo. Neste modelo foi possível fazer a ligação

de uma tabela com a outra de modo que a representação da interação existente entre os objetos envolvidos ficasse nítida.

**2.2. Desenvolvimento dos modelos lógico e físico**

O projeto lógico é processo de padronização de parâmetros de decisão sistêmica, são consideradas as fases anteriores do tratamento dos dados, passando-se então para a aplicabilidade (projeto físico) e tomada de decisão, sabendo-se os processos que culminaram à aplicação da lógica de decisão na resolução da problemática (ELMASRI e NAVATHE, 2011). O projeto denominado estação cavalo pantaneiro, sistema para auxiliar no processo de extração de informações e formulação de padrões, por meio de dados morfométricos da raça, para isto utilizando-se da árvore binária que trata se de um bloco de códigos algorítmicos que criam uma árvore de decisão e lógica *fuzzy*. A árvore da decisão é uma estrutura do BD que tem como fonte de relevância para classificação a lógica *fuzzy*, que tem como base de regras as variáveis contidas nas medidas lineares. Os dados de entrada são os valores de medidas lineares de morfometria, tendo como dados de saídas os índices corporais e de aptidão. Por se tratar de uma estrutura booleana 1 ou 0, que entra no *array* de repetição *while* (enquanto) for 1, do contrário o 0 sai do *array*, ou seja, sendo o 1 próximo do padrão ideal enquanto 0 é considerado distante. Para definir padrão das medidas lineares, considerou-se como valores ideais (classe desejável), as faixas de valores de animais referência da raça (intervalos de confiança); para a classe média, utilizou-se da mineração de dados considerando a faixa de valores de animais registrados com maior número de ocorrências e para a classe ruim com menor número de ocorrências. Para os índices morfométricos, foram considerados os índices existentes na literatura com padrões pré-estabelecidos para classificar os animais em longilíneos, mediolíneos e brevilíneos (McMANUS et al., 2008), consequentemente indicando a aptidão funcional dos cavalos (Tabela 2). O processo do teste de algoritmo minera os dados e classifica-os de acordo com as classes pré-definidas de morfometria (medidas lineares) e/ou aptidão conforme Tabela 2, cujo processo pode ser visualizado na Figura 1.

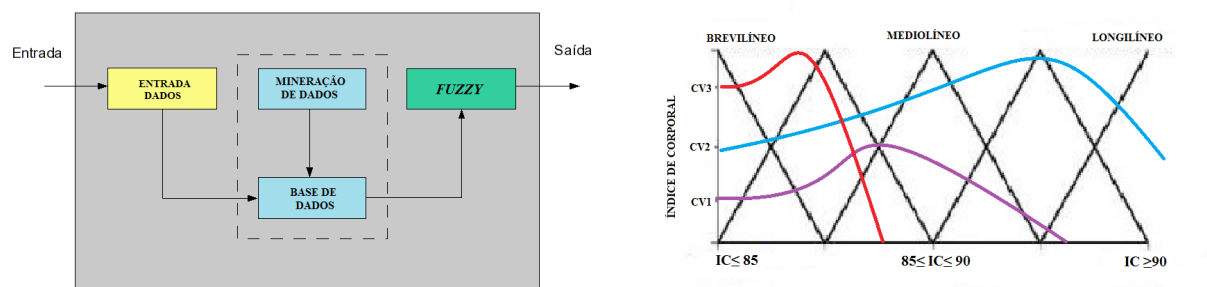


Figura 1. Regras para implementação da lógica *fuzzy* para se chegar aos padrões de conformação

Para o caso do desvio de padrão, em que as ocorrências são menores em comparação ao total, onde a média destacada (padrão da raça) e os resultados acontecem em intervalos incertos.

**3. Resultados e Discussão**

O SDECP por meio do gerenciamento do BD (HeidiSQL) extraiu pela mineração dos dados (WEKA) as 15 medidas morfométricas, para definir alguns padrões e relações da conformação animal. As 15 variáveis do sistema foram manipulados com auxílio matemático da lógica *fuzzy*, resultando na avaliação funcional da raça, por meio da análise da proximidade dos valores com a classificação pré-definida, como por exemplo, o IC que consta na Tabela 2.

Para especificar uma análise de índices conforme as classes descritas na Tabela 2, os intervalos resultantes da lógica *fuzzy* podem gerar outros dados “causa e efeito” (dados que geram outros dados), que na tomada de decisão *fuzzy* podem ser usados para avaliar outras medidas lineares e índices morfométricos descritos por SANTOS (2003).

Neste estudo, a aplicação da lógica *fuzzy* foi realizada para avaliar o índice corporal do cavalo Pantaneiro (Tabela 2), no qual considerou o desvio padrão próximo à média entre 85 e 90 como animal mediolíneo, que é o padrão desejado para a raça Pantaneira. Portanto, os demais valores serão considerados fora deste padrão. A lógica *fuzzy* permitiu classificar os animais de forma não pontual, e sim considerando as possíveis incertezas. Logo essas incertezas podem oscilar em poucos centímetros para mais ou para menos.

O mesmo procedimento poderá ser feito para outras medidas e outros índices morfométricos, tais como o índice de carga 1 e 2, índice torácico e índice dátilo torácico, contribuindo para o estudo de conformação e seleção funcional desta raça adaptada ao ecossistema Pantanal.

**4. Conclusões**

O sistema (SDECP) com suporte de um BD, associado com a mineração de dados foi capaz de informatizar e classificar os valores de medidas lineares e de índices morfométricos do cavalo Pantaneiro. Com a

implementação da lógica fuzzy foi possível classificar os animais em função da aptidão morfofuncional pré-estabelecida, contribuindo para a seleção e melhoramento da raça.

#### **Agradecimentos**

A Associação Brasileira de Criadores de Cavalos Pantaneiros (ABCCP) pelos dados fornecidos.

#### **Referências**

- ALVARENGA, L. Representação do conhecimento na perspectiva da ciência da informação em tempo e espaço digitais. Revista Eletrônica de Biblioteconomia e Ciência da Informação, Florianópolis, n. 15, jan.-jun. 2003.
- ELMASRI, R.; NAVATHE, S. B. Sistemas de Banco de Dados. Addison Wesley, 4a.edição, 2011.
- McMANUS, C.M.; SANTOS, S.A.; SILVA, J.A. da; LOUVANDINI, H.; ABREU, U.G.P.; RIBEIRO, D. B. O cavalo: raças, qualidades e defeitos. Rio de Janeiro: Globo, 1988. 290 p.
- SANTOS, S.A.; McMANUS, C.M.; MARIANTE, A.S. et al. Estratégias de conservação in situ do cavalo Pantaneiro. Corumbá: EMBRAPA, 2003. 29 p.
- SERENO, J.R.B.; ARIANTE, A.S. Body indices for the pantaneiro horse. Brazilian Journal of Veterinary Research and Animal Science, São Paulo, v. 45, n. 5, p. 362-370, 2008.
- WOLPP, R. L. G. Modelagem de uma ferramenta para construção de esquema de banco de dados usando o MER. Monografia do curso Especialização em Bancos de Dados do Instituto de Informática da Universidade Federal de Goiás, 2004.